

63548

**Delaval, Jan**

---

**From:** Gambel, Phillip  
**Sent:** Tuesday, April 02, 2002 9:15 AM  
**To:** Delaval, Jan  
**Subject:** exam answer 09 / 020393

jan

please perform a sequence and a sequence interference search for

ussn 09 / 020393

open and closed

amino acids 42-58 of SEQ ID NO: 3

thanx

phillip gambel  
art unit 1644  
308-3997

*Room 8 B03*

1644 mailbox

please let me know when it is ready. thanx.

**Jan Delaval**  
Reference Librarian  
Biotechnology & Chemical Library  
CM1 1E07 - 703-308-4498  
jan.delaval@uspto.gov

**THIS PAGE BLANK (USPTO)**

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval  
 Reference Librarian  
 Biotechnology & Chemical Library  
 CM1 1E07 - 703-308-4498  
 jan.delaval@uspto.gov

\*\*\*\*\*  
**STAFF USE ONLY****Type of Search****Vendors and cost where applicable**

|  |                          |                           |
|--|--------------------------|---------------------------|
| Searcher: <u>an</u>                    | NA Sequence (#) _____    | STN _____                 |
| Searcher Phone #: <u>4498</u>          | AA Sequence (#) <u>4</u> | Dialog _____              |
| Searcher Location: _____               | Structure (#) _____      | Questel/Orbit _____       |
| Date Searcher Picked Up: <u>7/2/02</u> | Bibliographic _____      | Dr.Link _____             |
| Date Completed: <u>7/2/02</u>          | Litigation _____         | Lexis/Nexis _____         |
| Searcher Prep & Review Time: _____     | Fulltext _____           | Sequence Systems <u>✓</u> |
| Clerical Prep Time: <u>15</u>          | Patent Family _____      | WWW/Internet _____        |
| Online Time: <u>20</u>                 | Other _____              | Other (specify) _____     |

**THIS PAGE BLANK (USPTO)**



63548  
**Delaval, Jan**

---

**From:** Gambel, Phillip  
**Sent:** Tuesday, April 02, 2002 10:07 AM  
**To:** Delaval, Jan  
**Subject:** second sequence to search for 09 /020393

jan

please perform a sequence and a sequence interference search

both open and closed

for

09 / 020393

SEQ ID NO: 18

thanx

phillip gambel  
art unit 1644  
308-3997

1644 mailbox 9E12

**Jan Delaval**  
**Reference Librarian**  
Biotechnology & Chemical Library  
CM1 1E07 – 703-308-4498  
jan.delaval@uspto.gov

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 12.82 seconds  
(without alignments)  
101.011 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVYTRRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : PIR\_68:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 96    | 100.0       | 128    | 1 RWHU59 | surface glycoprote  |
| 2          | 63    | 65.6        | 126    | 2 I36914 | CD59 protein - bab  |
| 3          | 63    | 65.6        | 128    | 2 I36894 | CD59 protein - gre  |
| 4          | 46.5  | 48.4        | 1302   | 2 T20767 | hypothetical prote  |
| 5          | 46.5  | 48.4        | 1767   | 2 T20766 | hypothetical prote  |
| 6          | 46    | 47.9        | 284    | 2 E83557 | probable transcrip  |
| 7          | 46    | 47.9        | 777    | 2 E83748 | hypothetical prote  |
| 8          | 45    | 46.9        | 1307   | 2 T21283 | hypothetical prote  |
| 9          | 44    | 45.8        | 205    | 2 H71639 | NADH dehydrogenase  |
| 10         | 44    | 45.8        | 807    | 2 D69102 | collagenase - Meth  |
| 11         | 43    | 44.8        | 544    | 2 S41626 | spike protein chain |
| 12         | 42    | 43.8        | 365    | 2 T20652 | hypothetical prote  |
| 13         | 42    | 43.8        | 419    | 2 C83681 | ABC transporter (s  |
| 14         | 42    | 43.8        | 464    | 2 T16889 | hypothetical prote  |
| 15         | 41    | 42.7        | 117    | 2 C69969 | hypothetical prote  |
| 16         | 41    | 42.7        | 179    | 2 S23358 | H+-transporting AT  |
| 17         | 41    | 42.7        | 223    | 1 VC8VCA | coat protein - tob  |
| 18         | 41    | 42.7        | 265    | 2 S64938 | hypothetical prote  |
| 19         | 41    | 42.7        | 410    | 2 T51212 | related to integr   |
| 20         | 41    | 42.7        | 720    | 2 T38647 | hypothetical prote  |
| 21         | 41    | 42.7        | 742    | 2 T23226 | hypothetical prote  |
| 22         | 41    | 42.7        | 805    | 2 T03896 | hypothetical prote  |
| 23         | 41    | 42.7        | 906    | 2 T48898 | disease resistance  |
| 24         | 41    | 42.7        | 908    | 2 T48898 | disease resistance  |
| 25         | 41    | 42.7        | 1162   | 2 S07421 | E2 glycoprotein pr  |
| 26         | 41    | 42.7        | 1162   | 2 S14939 | E2 glycoprotein pr  |
| 27         | 41    | 42.7        | 1162   | 2 S14940 | E2 glycoprotein pr  |
| 28         | 41    | 42.7        | 1528   | 2 S13743 | DNA strand transfe  |
| 29         | 41    | 42.7        | 1646   | 1 WMTMS2 | 186k protein - cuc  |

|    |      |      |      |          |                     |
|----|------|------|------|----------|---------------------|
| 30 | 41   | 42.7 | 3433 | 1 S28381 | utrophin - human    |
| 31 | 40   | 41.7 | 78   | 2 H84010 | hypothetical prote  |
| 32 | 40   | 41.7 | 165  | 2 F71691 | hypothetical prote  |
| 33 | 40   | 41.7 | 218  | 2 F69972 | probable membrane   |
| 34 | 40   | 41.7 | 244  | 2 T00825 | probable heat choc  |
| 35 | 40   | 41.7 | 497  | 2 T06727 | hypothetical prote  |
| 36 | 40   | 41.7 | 607  | 2 T43322 | meiosis specific p  |
| 37 | 40   | 41.7 | 742  | 1 A43344 | synaptic vesicle p  |
| 38 | 40   | 41.7 | 742  | 2 S27263 | synaptic vesicle p  |
| 39 | 39.5 | 41.1 | 156  | 2 T31664 | DAP-kinase homolog  |
| 40 | 39.5 | 41.1 | 169  | 2 B64374 | hypothetical prote  |
| 41 | 39.5 | 41.1 | 632  | 2 T31667 | guanylate cyclase   |
| 42 | 39.5 | 41.1 | 632  | 2 T31666 | guanylate cyclase   |
| 43 | 39.5 | 41.1 | 760  | 2 S75517 | natrurietic peptid  |
| 44 | 39.5 | 41.1 | 854  | 1 VCLJST | env polyprotein pr  |
| 45 | 39   | 40.6 | 119  | 2 T32569 | hypothetical prote  |
| 46 | 39   | 40.6 | 258  | 2 T13591 | tail fiber adhesin  |
| 47 | 39   | 40.6 | 289  | 2 S73935 | hypothetical prote  |
| 48 | 39   | 40.6 | 287  | 2 I38517 | synthaxin - human   |
| 49 | 39   | 40.6 | 297  | 2 S52726 | synthaxin-4 - human |
| 50 | 39   | 40.6 | 332  | 2 C96693 | hypothetical prote  |

## ALIGNMENTS

RESULT 1  
RWHU59  
surface glycoprotein CD59 precursor [validated] - human  
N/Alternate names: 1F5 antigen protein; 20K homologous restriction factor (HRF20); CD  
plex inhibition factor (MACIF); membrane inhibitor of reactive lysis (MIRL); protecti  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1990 #sequence, revision 30-Sep-1990 #text, change 08-Dec-2000  
C/Accession: A46252; J010109; A33405; J010134; A34587; S05504; S09201; A60828; PL0041;  
R/Betranka, J.G.; Fleenor, D.E.; Sykes, K.; Kaufman, R.E.; Rosse, W.F.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7876-7879, 1992  
A/Title: Structure of the CD59-encoding gene: Further evidence of a relationship to m  
A/Reference number: A46252; M01D:92390353  
A/Accession: A46252  
A/Molecule type: DNA  
A/Residues: 1-128 <PEP>  
A/Cross-references: GB:M82840; NID:9180149; PIDN:AAA68793.1; PID:9180151  
A/Note: sequence extracted from NCBI backbone (NCBIN:112714, NCBIN:112718, NCBIN:1127  
R/Davies, A.; Simmons, D.L.; Hale, G.; Harrison, R.A.; Tighe, H.; Lachmann, F.J.; Mal  
J. Exp. Med. 170, 637-654, 1989  
A/Title: CD59, an LY-6-like protein expressed in human lymphoid cells, regulates the  
A/Reference number: J010109; M01D:89361238  
A/Accession: J010109  
A/Molecule type: mRNA  
A/Residues: 1-128 <DAV>  
A/Cross-references: EMBL:X16447; NID:929805; PIDN:CAA34467.1; PID:929806  
R/Okada, H.; Nagami, Y.; Takahashi, K.; Okada, N.; Hideshima, T.; Takizawa, H.; Kondo  
Biochem. Biophys. Res. Commun. 167, 1553-1559, 1989  
A/Title: 20 kDa homologous restriction factor of complement resembles T cell activati  
A/Reference number: A33405; M01D:89350963  
A/Accession: A33405  
A/Molecule type: mRNA  
A/Residues: 1-128 <OKA>  
A/Cross-references: GB:M27909; NID:9623406; PIDN:AAA60543.1; PID:9623407  
R/Sugita, Y.; Tobe, T.; Oda, E.; Tomita, M.; Yasukawa, K.; Yamaji, N.; Takemoto, T.;  
J. Biochem. 106, 555-557, 1989  
A/Title: Molecular cloning and characterization of MACIF, an inhibitor of membrane ch  
A/Reference number: J010134; M01D:90110046  
A/Accession: J010134  
A/Molecule type: mRNA  
A/Residues: 1-128 <SUG>  
A/Note: parts of this sequence, including the amino end of the mature protein, were c  
A/Note: sites for glycosylation and the absence of glycosylation were confirmed  
R/Sawada, R.; Ohashi, K.; Anaguchi, H.; Okazaki, H.; Hattori, M.; Minato, N.; Naruto,  
DNA Cell Biol. 9, 213-220, 1990  
A/Title: Isolation and expression of the full-length cDNA encoding CD59 antigen of hu  
A/Reference number: A34587; M01D:90253615  
A/Accession: A34587

A: Molecule type: mRNA  
 A: Residues: 1-128 <SAM>  
 A: Cross-references: GB:M34671; NID:q180152; PIDN:AAA51952.1; PID:q180153  
 R: Sawada, R.; Ohashi, K.; Okano, K.; Hattori, M.; Minato, N.; Naruto, M.  
 A: Title: Complementary DNA sequence and deduced peptide sequence for CD59/MEW-43 antigen  
 A: Reference number: S05504; MUID:89386002  
 A: Accession: S05504  
 A: Molecule type: mRNA  
 A: Residues: 27-128 <SA2>  
 A: Cross-references: EMBL:X15861; NID:q29803; PIDN:CAA33870.1; PID:q1340180  
 R: Philbrick, W.M.; Palfree, R.G.E.; Maher, S.E.; Bridgett, M.M.; Sirlin, S.; Bothwell, A.  
 A: Title: The CD59 antigen is a structural homologue of murine Ly-6 antigens but lacks its  
 A: Reference number: S09201; MUID:90168959  
 A: Accession: S09201  
 A: Molecule type: mRNA  
 A: Residues: 1-128 <PHD>  
 A: Cross-references: EMBL:X17198; NID:q29814; PIDN:CAA35059.1; PID:q29815  
 R: Cabral, A.R.; Cole, L.A.; Walz, D.A.; Castor, C.W.  
 A: Title: Connective tissue activation, XXXII. Structural and biologic characteristics of  
 A: Reference number: A60828; MUID:88134429  
 A: Accession: A60828  
 A: Molecule type: protein  
 A: Residues: 26-27, 'V', 29-30, 'D', 32-37, 'X', 39-42, 'XX', 45-50, 'X', 52-62; 'VXRLID' <CAB>  
 A: Experimental source: normal urine  
 A: Note: the six unknown or mismatched residues in the amino-terminal fragment correspond  
 R: Stefanova, I.; Hilgert, I.; Kristofova, H.; Brown, R.; Low, M.G.; Horejsl, V.  
 A: Title: Characterization of a broadly expressed human leucocyte surface antigen MEW-43  
 A: Reference number: P10041; MUID:89143489  
 A: Accession: P10041  
 A: Molecule type: protein  
 A: Residues: 26-42 <SRP>  
 R: Harada, R.; Okada, N.; Fujita, T.; Okada, H.  
 A: Title: Purification of IFS antigen that prevents complement attack on homologous cell  
 A: Reference number: A60774; MUID:90171576  
 A: Accession: A60774  
 A: Molecule type: protein  
 A: Residues: 26-42, 'XX', 45-50, 'X', 52, 'X', 54-57, 'X', 59-63 <HAR>  
 R: Nishimura, H.; Stewart, B.H.; Rollins, S.A.; Zhao, J.; Bothwell, A.L.M.; Sims, P.J.  
 A: Title: Contribution of the N-linked carbohydrate of erythrocyte antigen CD59 to its CD  
 A: Reference number: A38089; MUID:92235065  
 A: Accession: A38089  
 A: Molecule type: protein  
 A: Residues: 40-42, 'X', 44-49 <NIN>  
 R: Sugita, Y.; Nakano, Y.; Oda, E.; Noda, K.; Tobe, T.; Miura, N.H.; Tomita, M.  
 A: Title: Determination of carboxyl-terminal residue and disulfide bonds of MAC1F(CD59),  
 A: Reference number: PX0068; MUID:94103166  
 A: Accession: PX0068  
 A: Molecule type: protein  
 A: Residues: 26-29, 30-39, 40-42, 'X', 44-48, 49-52, 56-65, 65-72, 88-90, 92-96 <SD2>  
 R: Tone, M.; Walsh, L.A.; Waldmann, H.  
 A: Title: Gene structure of human CD59 and demonstration that discrete mRNAs are generated  
 A: Reference number: I37223; MUID:93021133  
 A: Accession: I37223  
 A: Status: translated from GB/EMBL/DBJ  
 A: Molecule type: DNA  
 A: Residues: 1-128 <RES>  
 A: Cross-references: EMBL:Z14113; NID:q29810; PIDN:CAA78486.1; PID:q825637  
 C: Comment: This cell surface antigenic glycoprotein inhibits homologous complement lysis  
 C: Comment: of the binding of C9 and C8 to C5b  
 C: Comment: This glycosylphosphatidylinositol-anchored protein is deficient in cases of B  
 A: Genes: GDB:CD59  
 A: Cross-references: GDB:119769; OMIM:107271  
 A: Map position: 11p13-11p13

A: Introns: 23/1: 57/1  
 A: Note: The first intron occurs before the initiator codon  
 C: Species: Papio sp. (baboon)  
 C: Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein; phosph  
 F: 1-25/Domain: signal sequence #status predicted: lipoprotein; membrane protein; phosph  
 F: 26-102/Product: surface glycoprotein CD59 #status experimental <SIG>  
 F: 103-128/Domain: Ly-6 homology <LY6>  
 F: 128-51, 31-38, 44-64/Disulfide bonds: #status experimental <MAT>  
 F: 33/Binding site: carboxyl-terminal propenyl #status predicted <CTP>  
 F: 43/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F: 70-88, 89-94/Disulfide bonds: (or 70-89, 88-94) #status experimental  
 F: 102/Modified site: GPI anchor ethanolamine amidated carboxyl end (Asn) (in mature 1

## Query Match

Best Local Similarity 100.0%; Score 96; DB 1; Length 128;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRLRNE 17  
 |||||||||||||

DB 67 FANCFNDISTLRKE 83

## RESULT 2

CD59 protein - baboon

C: Species: Papio sp. (baboon)

C: Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C: Accession: I36914

R: Fodor, W.L.; Rollins, S.A.; Bianco-Caron, S.; Burton, W.V.; Guilmette, E.R.; Rother

A: Title: Primate terminal complement inhibitor homologues of human CD59.

A: Reference number: I36894; MUID:95104908

A: Accession: I36914

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-126 <RES>

A: Cross-references: GB:I22862; NID:q514327; PIDN:AAA74127.1; PID:q514328

C: Superfamily: Ly-6 antigen; Ly-6 homology

F: 26-100/Domain: Ly-6 homology <LY6>

## Query Match

Best Local Similarity 65.6%; Score 63; DB 2; Length 126;  
 Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRLRNE 17  
 |:|||||:|:|:|

DB 67 FANCFNDISTLRKE 83

## RESULT 3

CD59 protein - green monkey

C: Species: Cercopithecus aethiops (green monkey, grivet)

C: Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C: Accession: I36894

R: Fodor, W.L.; Rollins, S.A.; Bianco-Caron, S.; Burton, W.V.; Guilmette, E.R.; Rother

A: Title: Primate terminal complement inhibitor homologues of human CD59.

A: Reference number: I36894; MUID:95104908

A: Accession: I36894

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-128 <RES>

A: Cross-references: GB:I22863; NID:q514314; PIDN:AAA74126.1; PID:q514315

C: Superfamily: Ly-6 antigen; Ly-6 homology

F: 26-102/Domain: Ly-6 homology <LY6>

## Query Match

Best Local Similarity 58.8%; Score 63; DB 2; Length 128;  
 Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

|    | Matches | 10; Conservative   | 5; Mismatches | 2; Indels | 0; Gaps |
|----|---------|--------------------|---------------|-----------|---------|
| Qy | 1       | FEHCNFNDVYTRLRENE  | 17            |           |         |
|    | 1       | : : : : : :        | : : : : : :   |           |         |
| Db | 67      | FANCFNFDISTILKRESE | 83            |           |         |

RESULT 4  
T20767  
hypothetical protein Fl1C1.5b - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T20767  
R:Palmer, S.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: Z19321  
A:Accession: T20767  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1302 <WII>  
A:Cross-references: EMBL:554270; PIDN:CAA91031.1; GSPDB:GN00028; CESP:Fl1C1.5b  
A:Experimental source: clone Fl1C1  
C:Genetics:  
A:Gene: CESP:Fl1C1.5b  
A:Map position: X  
A:Introns: 21/1, 97/3, 134/3, 189/3, 245/3, 284/3, 474/2, 553/3, 701/2, 734/3, 807/1, 88

RESULT 5  
T20766  
hypothetical protein F11C1.5a - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T20766  
R:Palmer, S.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: Z19321  
A:Accession: T20766  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1767 <MIL>  
A:Cross-references: EMBL:Z54270; PIDN:CAA91030.1; GSPDB:GN00028; CESP:F11C1.5a  
A:Experimental source: clone F11C1  
C:Genetics:  
A:Gene: CESP:F11C1.5a  
A:Map position: X  
A:Introns: 21/1: 97/3; 134/3; 189/3; 245/3; 284/3; 474/2; 553/3; 701/2; 734/3; 807/1; 889/3; 1658/3; 1700/2

```
RESULT      6
E83557
probable transcription regulator PA0708 [imported] - Pseudomonas aeruginosa (strain PA01)
c:Species: Pseudomonas aeruginosa
c:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
```

R.Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardblyg, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A.Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa  
A.Reference number: AB29550; MUID:20437337  
A.Accession: E83557  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-284 <STO>  
A.Cross-references: GB:AE004506; GB:AE004091; NID:99946584; PIDN:AMG04097.1; GSPDB:GN  
C.Experimental source: strain PA01  
C.Genetics:  
A.Gene: PA0708

```

RESULT      7
EB3748
hypothetical protein BH0789 [imported] - Bacillus halodurans (strain C-125)
C|Species: Bacillus halodurans
C|Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C|Accession: EB3748
R|Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A|Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A|Reference number: AB3650; MUID:20263314
A|Accession: EB3748
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1777 <SNO>
A|Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BAB04508.1; GSPDB:G
A|Experimental source: strain C-125
C|Genetics:
A|Gene: BH0789

```

```

RESULT      8
T121283
hypothetical protein F23A7.5 - Caenorhabditis elegans
C1:Species: Caenorhabditis elegans
C1:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C1:Accession: T121283
R:McMurray, A.
  submitted to the EMBL Data Library, October 1996
A1:Reference number: Z19401
A1:Accession: T121283
A1:Status: preliminary; translated from GR/EMBL/DBDJ
A1:Molecule type: DNA
A1:Residues: 1-1307 <NTL>
A1:Cross-references: EMBL:Z81067; P1DN:CA802976.1; GSPDB:GM00028; CESP:F23A7.5
A1:Experimental source: Clone F23A7
C1:Genetics:
A1:Gene: CESP:F23A7.5
A1:Map position: X

```

A; Introns: 154/1; 230/2; 271/1; 307/1; 466/2; 551/2; 627/3; 663/1; 699/3; 746/3; 772/3;

## Query Match

Best Local Similarity 46.9%; Score 45; DB 2; Length 1307;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 654 CKNNDVTTLREN 16  
| | | | |  
| | | | |

## RESULT

9

H71639

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain J RP790 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
R:Accession: H71639  
R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alismark, U.  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499  
A:Accession: H71639  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-205 <AND>  
A:Cross-references: GB:AJ235273; GB:AJ235269; NID:93861237; PIDN:CA15216.1; PID:9386131  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: nuoJ; RP790  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 6  
C:Keywords: membrane-associated complex; NAD; oxidoreductase

## Query Match

Best Local Similarity 66.7%; Score 44; DB 2; Length 205;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 77 NFNDVTRLREN 16  
| | | | |  
| | | | |

## RESULT

10

D69102

collagenase - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Kl, S.; Church, G.M.; Daniels, C.D.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7151, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514  
A:Accession: D69102  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-807 <MTH>  
A:Cross-references: GB:AE000931; GB:AE000666; NID:g2622885; PIDN:AA86229.1; PID:g262289  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1763  
A:Start codon: TTG

## Query Match

Best Local Similarity 45.8%; Score 44; DB 2; Length 807;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 562 FEHCNNDVTTLRE 15  
| | | | |  
| | | | |

## RESULT

11

S41626

Spike protein chain 1 precursor - avian infectious bronchitis virus (fragment)  
C:Species: avian infectious bronchitis virus, IBV  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 20-Sep-1999  
R:Accession: S41626  
R:Canavan, D.; Davis, P.J.  
Arch. Virol. 130, 471-476, 1992  
A:Title: Sequence analysis of strains of avian infectious bronchitis coronavirus isol  
A:Reference number: S41626  
A:Accession: S41626  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-544 <CAV>  
A:Cross-references: EMBL:X64737; NID:9453158; PIDN:CAA46003.1; PID:9453159  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1992  
C:Superfamily: coronavirus E2 glycoprotein

## Query Match

Best Local Similarity 75.0%; Score 43; DB 2; Length 544;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 101 HCNETDIT 10  
| | | | |  
| | | | |

## RESULT

12

T20652

hypothetical protein Y102A5C.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
R:Morimoto, B.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: T20652  
A:Accession: T20652  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-365 <NIT>  
A:Cross-references: EMBL:Z81496; PIDN:CAB04075.1; GSPDB:GN00023; CESP:Y102A5C.1  
R:Gardner, A.  
submitted to the EMBL Data Library, September 1998  
A:Reference number: Z20204  
A:Accession: Z20204  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-365 <NIT>  
A:Cross-references: EMBL:AL031627; PIDN:CAA20972.1; CESP:Y102A5C.1  
A:Experimental source: clone Y102A5C  
C:Genetics:  
A:Gene: CESP:Y102A5C.1  
A:Map position: 5  
A:Introns: 93/2; 104/1; 298/3

## Query Match

Best Local Similarity 43.8%; Score 42; DB 2; Length 365;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Db 288 FEHLCFQDVKERLES 303  
| | | | |  
| | | | |

## RESULT

13

C83681

ABC transporter (substrate-binding protein) BH0251 [Imported] - Bacillus halodurans (



Db 3 NENOITFLAENE 15

RESULT 17

VCBVC

C:Species: tobacco rattie virus (strain CAM)

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991

C:Accession: A26027

R:Bergh, S.T.; Koziel, M.G.; Huang, S.C.; Thomas, R.A.; Gilley, D.P.; Siegel, A.

Nucleic Acids Res. 13, 8507-8518, 1985

A:Reference number: A26027; MUID:86093657

A:Accession: A26027

A:Molecule type: genomic RNA

A:Residues: 1-723 <BER>

A:Cross-references: GB:X03241; NID:962057; PIDN:CAA26998.1; PID:662058

C:Genetics: the authors translated the codon GAG for residue 178 as Gly

A:Map position: segment 2

C:Superfamily: tobacco rattie virus coat protein

C:Keywords: coat protein

Query Match

Best Local Similarity 42.7%; Score 41; DB 1; Length 223;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NFNVDVTRLR 14

Db 25 NMRKDVTRLR 34

RESULT 18

S64938

hypothetical protein YLR102c - Yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 01-Aug-1995 #sequence\_revision 24-May-1996

C:Accession: S64938

R:Messenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64935

A:Accession: S64938

A:Molecule type: DNA

A:Residues: 1-265 <MESS>

A:Cross-references: EMBL:Z73274; NID:q1360485; PIDN:CAA97667.1; PID:q1360486; GSPDB:GN00

C:Genetics: experimental source: strain S288C

A:Gene: MIPS:YLR102c

A:Map position: 12R

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YLR102c

Query Match

Best Local Similarity 42.7%; Score 41; DB 2; Length 265;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRLR 15

Db 26 FNKAFNNFTPLRK 40

RESULT 19

T51212

related to integral membrane protein pth11 [imported] - Neurospora crassa

N:Alternate names: protein B24M22.80

C:Species: Neurospora crassa

C:Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000

C:Accession: T51212

R:Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25286

A:Accession: T51212

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-410 <SCH>

A:Cross-references: EMBL:AL390354; GSPDB:GN00116; NCSP:B24M22.80

A:Experimental source: strain OR74A

C:Genetics:

A:Gene: NCSP:B24M22.80

A:Map position: 6

A:Introns: 77/3; 251/2

Query Match

Best Local Similarity 42.7%; Score 41; DB 2; Length 410;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 CNFNDVTRLRNE 17

Db 357 CTYNDKSRNRND 370

RESULT 20

T38647

hypothetical protein SPAC32A11.01 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999

C:Accession: T38647

R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1996

A:Reference number: T21803

A:Accession: T38647

A:Molecule type: DNA

A:Residues: 1-720 <SKB>

A:Cross-references: EMBL:Z69796; PIDN:CAA93698.1; GSPDB:GN00066; SPDB:SPAC32A11.01

C:Genetics: experimental source: strain 972h-; cosmid c32A11

A:Gene: SPDB:SPAC32A11.01

A:Map position: 1

A:Introns: 45/1; 210/2

C:Superfamily: Schizosaccharomyces hypothetical protein SPAC32A11.01

Query Match

Best Local Similarity 42.7%; Score 41; DB 2; Length 720;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNFDVTR 11

Db 383 FSHCSFESLTT 393

RESULT 21

T23226

hypothetical protein K02B7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999

C:Accession: T23226

R:Percy, C.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19711

A:Accession: T23226

A>Status: Preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-742 <WIL>

A:Cross-references: EMBL:Z83112; PIDN:CAB05538.1; GSPDB:GN00020; CESP:K02B7.1

A:Experimental source: clone K02B7

C:Genetics:

A:Gene: CESP:K02B7.1

A:Map position: 2

A:Introns: 31/1; 196/2; 281/2; 357/3; 393/1; 429/3; 474/2; 496/3; 546/2

C:Superfamily: Caenorhabditis elegans hypothetical protein W03G1.4



Query Match 42.7%; Score 41; DB 2; Length 742;  
Best Local Similarity 61.5%; Pred. No. 98;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CKNFNDVTRLREN 16  
| | | | |  
Db 384 CKINDEYTLREN 396

## RESULT 22

T03896  
hypothetical protein R09E12.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 01-Dec-2000  
C:Accession: T03896; T32395; T33999  
R:Blanchard, M.; Elliott, G.; Keppler, D.; Smith, A.  
submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of *C. elegans* cosmid R09E12.  
A:Reference number: Z15130  
A:Accession: T03896  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-805 <BLA>  
A:Cross-references: EMBL:AF016422; NID:g2291187; PIDN:AB65317.1; PID:g2291193  
R:Magri, L.; Scheet, P.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of *C. elegans* cosmid K02E7.  
A:Reference number: Z21159  
A:Accession: T32395  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-281, 'H', 283-805 <MAG>  
A:Cross-references: EMBL:AF025465; PIDN:AA071023.1; GSPDB:GN00020; CESP:K02E7.3  
A:Experimental source: Strain Bristol N2; clone K02E7  
R:Paulley, A.; Scheet, P.; Harper, M.  
submitted to the EMBL Data Library, February 1999  
A:Description: The sequence of *C. elegans* cosmid W03G1.  
A:Reference number: Z21454  
A:Accession: T33999  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-504, 'S', 506-805 <PAU>  
A:Cross-references: EMBL:AF125964; PIDN:AD14752.1; GSPDB:GN00022; CESP:W03G1.4  
A:Experimental source: strain Bristol N2; clone W03G1  
C:Genetics:  
A:Gene: CESP:K02E7.3; CESP:W03G1.4  
A:Map position: V  
A:Introns: 31/1; 196/2; 281/2; 357/3; 393/1; 429/3; 489/3; 533/3; 559/3; 609/2  
A>Note: R09E12.6  
C:Superfamily: *Caenorhabditis elegans* hypothetical protein W03G1.4

Query Match 42.7%; Score 41; DB 2; Length 805;  
Best Local Similarity 61.5%; Pred. No. 1.1e+02;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CKNFNDVTRLREN 16  
| | | | |  
Db 384 CKINDEYTLREN 396

## RESULT 23

T48898  
disease resistance protein RPP8 [validated] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 21-Jul-2000  
C:Accession: T48898  
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.; Dangl  
Plant Cell 10, 1861-1874, 1998  
A:Title: Intragenic recombination and diversifying selection contribute to the evolution  
A:Reference number: Z24999; MUID:99030193  
A:Accession: T48898

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-906 <MCD>  
A:Cross-references: EMBL:AF089710; NID:g3928861; PIDN:AC83165.1; PID:g3928862  
A:Experimental source: Landsberg erecta  
C:Genetics:  
A:Gene: RPP8  
A:Introns: 293/1; 342/1  
A:Function:  
A:Description: promotes resistance to *Peronospora parasitica*

QY 2 EHCNFDVTRLRE 15  
| | | | |  
Db 703 ERCNFETLSLRE 716

Query Match 42.7%; Score 41; DB 2; Length 906;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNFDVTRLRE 15  
| | | | |  
Db 703 ERCNFETLSLRE 716

## RESULT 24

T48899  
disease resistance protein rpp8 [similarity] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 21-Jul-2000  
C:Accession: T48899  
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.; Da  
Plant Cell 10, 1861-1874, 1998  
A:Title: Intragenic recombination and diversifying selection contribute to the evol  
A:Reference number: Z24999; MUID:99030193  
A:Accession: T48899  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-908 <MCD>  
A:Cross-references: EMBL:AF089711; NID:g3901293; PIDN:AC78631.1; PID:g3901294  
A:Experimental source: Columbia  
C:Genetics:  
A:Gene: rpp8  
A:Introns: 293/1; 342/1  
A:Function:  
A:Description: susceptible allele of a gene that promotes resistance to *Peronospora p*

Query Match 42.7%; Score 41; DB 2; Length 908;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNFDVTRLRE 15  
| | | | |  
Db 705 ERCNFETLSLRE 718

## RESULT 25

S07421  
E2 glycoprotein precursor - avian infectious bronchitis virus  
N:Alternate names: peplomer glycoprotein  
N:Contains: E2 glycoprotein subunit S1; E2 glycoprotein subunit S2  
C:Species: avian infectious bronchitis virus, IBV  
C>Date: 18-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Sep-1999  
C:Accession: S07421  
R:Niesters, H.G.M.; Lenstra, J.A.; Spaan, W.J.M.; Zijderfeld, A.J.; Bleumink-Pluym, N  
Virus Res. 5, 253-263, 1986  
A:Title: The peplomer protein sequence of the M41 strain of coronavirus IBV and its c  
A:Reference number: S07421; MUID:87021475  
A:Accession: S07421  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1162 <NIE>  
A:Cross-references: EMBL:M21883; NID:g331183; PIDN:AA66575.1; PID:g331184  
C:Superfamily: coronavirus E2 glycoprotein  
C:Keywords: glycoprotein

Tue Apr 2 09:51:57 2002

us-09-020-393b-3\_copy\_42\_58.open.rpr

Query Match  
Best Local Similarity 42.7%; Score 41; DB 2; Length 1162;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 HCNFNDVT 10  
|||:|  
Db 101 HCNFSDT 108

Search completed: April 2, 2002, 09:20:13  
Job time: 22 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:20:16 ; Search time 10.15 seconds  
(without alignments)  
61.409 Million cell updates/sec

Title: US-09-020-393B-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 96    | 100.0       | 128    | 1     | CD59_HUMAN  |
| 2          | 63    | 65.6        | 126    | 1     | CD59_PAPSP  |
| 3          | 63    | 65.6        | 128    | 1     | CD59_CERAE  |
| 4          | 49    | 51.0        | 124    | 1     | CD59_RABIT  |
| 5          | 49    | 51.0        | 128    | 1     | CD59_AOTTR  |
| 6          | 47    | 49.0        | 128    | 1     | CD59_CALSO  |
| 7          | 42    | 43.8        | 464    | 1     | Y5V5_CAEEL  |
| 8          | 41    | 42.7        | 179    | 1     | ATPF_ODOSI  |
| 9          | 41    | 42.7        | 223    | 1     | COAT_TKPCA  |
| 10         | 41    | 42.7        | 720    | 1     | YD71_SCHPO  |
| 11         | 41    | 42.7        | 1162   | 1     | YGL2_IBVB   |
| 12         | 41    | 42.7        | 1162   | 1     | YGL2_IBVM   |
| 13         | 41    | 42.7        | 1528   | 1     | KEM1_YEAST  |
| 14         | 41    | 42.7        | 1648   | 1     | RKPO_CGMVS  |
| 15         | 41    | 42.7        | 3433   | 1     | UTRO_HUMAN  |
| 16         | 40    | 41.7        | 165    | 1     | RIMM_RICPR  |
| 17         | 40    | 41.7        | 212    | 1     | HFA3_HAEIN  |
| 18         | 40    | 41.7        | 231    | 1     | FLPA_SULAC  |
| 19         | 40    | 41.7        | 607    | 1     | KMS1_SCHPO  |
| 20         | 40    | 41.7        | 716    | 1     | HSB3_DROAV  |
| 21         | 40    | 41.7        | 742    | 1     | SYV2_RAT    |
| 22         | 39.5  | 41.1        | 169    | 1     | Y594_METJA  |
| 23         | 39.5  | 41.1        | 760    | 1     | SPOT_SYNY3  |
| 24         | 39.5  | 41.1        | 854    | 1     | ENV_STYVZ   |
| 25         | 39    | 40.6        | 123    | 1     | CD59_PIG    |
| 26         | 39    | 40.6        | 224    | 1     | PVA_PLAFA   |
| 27         | 39    | 40.6        | 289    | 1     | Y084_MYCPN  |
| 28         | 39    | 40.6        | 297    | 1     | STX4_HUMAN  |
| 29         | 39    | 40.6        | 353    | 1     | YC54_HAEIN  |
| 30         | 39    | 40.6        | 574    | 1     | VB18_VACC   |
| 31         | 39    | 40.6        | 574    | 1     | VB18_VACCV  |
| 32         | 39    | 40.6        | 688    | 1     | YJ80_YEAST  |
| 33         | 39    | 40.6        | 1678   | 1     | CA64_HUMAN  |

## ALIGNMENTS

| RESULT | 1   | CD59_HUMAN | STANDARD; | PRT; | 128 AA. |
|--------|---|------------|-----------|------|---------|
| ID     | CD59_HUMAN  |            |           |      |         |
| AC     | P13987;   |            |           |      |         |
| DT     | 01-JAN-1990 (Rel. 13, Created)  |            |           |      |         |
| DT     | 01-JAN-1990 (Rel. 13, Last sequence update)   |            |           |      |         |
| DT     | 20-AUG-2001 (Rel. 40, Last annotation update)   |            |           |      |         |
| DE     | CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (MEM43 ANTIGEN) (PROTECTIN) (MEMBRANE INHIBITOR OF REACTIVE LYSIS) (MIRL) (20 KDA DE HOMOLOGOUS RESTRICTION FACTOR) (HRF-20) (HRF20) (1F5 ANTIGEN). |            |           |      |         |
| GN     | CD59.   |            |           |      |         |
| OS     | Homo sapiens (Human).   |            |           |      |         |
| OC     | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;   |            |           |      |         |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |            |           |      |         |
| OX     | NCBI_TaxID=9606;  |            |           |      |         |
| RN     | [1]   |            |           |      |         |
| RP     | SEQUENCE FROM N.A.  |            |           |      |         |
| RC     | TISSUE-T-cell;  |            |           |      |         |
| RX     | MEDLINE=89361238; PubMed=2475570;   |            |           |      |         |
| RA     | Davies A., Simmons D.L., Hale G., Harrison R.A., Tishe H.,  |            |           |      |         |
| RA     | Lachmann P.J., Waldmann H.;   |            |           |      |         |
| RT     | "CD59, an ly-6-like protein expressed in human lymphoid cells,  |            |           |      |         |
| RT     | regulates the action of the complement membrane attack complex on   |            |           |      |         |
| RT     | homologous cells.";   |            |           |      |         |
| RL     | J. Exp. Med. 170:637-654(1989).   |            |           |      |         |
| RN     | [2]   |            |           |      |         |
| RP     | SEQUENCE FROM N.A.  |            |           |      |         |
| RX     | MEDLINE=90168959; PubMed=1689664;   |            |           |      |         |
| RA     | Philbrick W.M., Palfrey R.G.E., Roger G.E., Maher S.E.,   |            |           |      |         |
| RT     | "The CD59 antigen is a structural homologue of murine Ly-6 antigens   |            |           |      |         |
| RT     | but lacks interferon inducibility.";  |            |           |      |         |
| RL     | Eur. J. Immunol. 20:87-92(1990).  |            |           |      |         |
| RN     | [3]   |            |           |      |         |
| RP     | SEQUENCE FROM N.A.  |            |           |      |         |
| RX     | MEDLINE=89350983; PubMed=2475111;   |            |           |      |         |
| RA     | Okada H., Nagami Y., Takahashi K., Okada N., Hideshima T.,  |            |           |      |         |
| RA     | Takikawa H., Kondo J.;  |            |           |      |         |
| RT     | "20 KDa homologous restriction factor of complement resembles T cell  |            |           |      |         |
| RT     | activating protein.";   |            |           |      |         |
| RL     | Biochem. Biophys. Res. Commun. 162:1553-1559(1989).   |            |           |      |         |
| RN     | [4]   |            |           |      |         |
| RP     | SEQUENCE FROM N.A.  |            |           |      |         |
| RX     | MEDLINE=90110046; PubMed=2606909;   |            |           |      |         |
| RA     | Sugita Y., Tobe T., Oda E., Tomita M., Yasukawa K., Yamaji N.,  |            |           |      |         |
| RA     | Takemoto T., Furuchi K., Takayama M., Yano S.;  |            |           |      |         |
| RT     | "Molecular cloning and characterization of MACIF, an inhibitor of   |            |           |      |         |
| RT     | membrane channel formation of complement.";   |            |           |      |         |
| RL     | J. Biochem. 106:555-557(1989).  |            |           |      |         |
| RN     | [5]   |            |           |      |         |

|    |      |      |      |   |            |                    |
|----|------|------|------|---|------------|--------------------|
| 34 | 39   | 40.6 | 1888 | 1 | YD72_SCHPO | 014207 schizosacch |
| 35 | 38   | 39.6 | 344  | 1 | YD76_SCHPO | 014220 schizosacch |
| 36 | 38   | 39.6 | 502  | 1 | KDC2_DROME | P16912 drosophila  |
| 37 | 38   | 39.6 | 1233 | 1 | HCV4_OCTDO | P12659 octopus dof |
| 38 | 38   | 39.6 | 1450 | 1 | RP01_ASEB7 | P42486 african swi |
| 39 | 38   | 39.6 | 2896 | 1 | HCV6_OCTDO | 061363 octopus dof |
| 40 | 37.5 | 39.1 | 517  | 1 | Y4CC_RHISN | P53385 rhizobium s |
| 41 | 37   | 38.5 | 137  | 1 | COPI_DICDI | P54706 dicystosell |
| 42 | 37   | 38.5 | 273  | 1 | RL2_RICPR  | 092c98 rickettsia  |
| 43 | 37   | 38.5 | 420  | 1 | PEL_BACSU  | P39116 bacillus su |
| 44 | 37   | 38.5 | 516  | 1 | SMP3_YEAST | 004174 saccharomyc |
| 45 | 37   | 38.5 | 535  | 1 | ATP2_YEAST | P53296 saccharomyc |
| 46 | 37   | 38.5 | 684  | 1 | YMW5_YEAST | 004779 saccharomyc |
| 47 | 37   | 38.5 | 758  | 1 | PURL_CORAM | 091hw9 corynebacte |
| 48 | 37   | 38.5 | 853  | 1 | VMTH_LAMBD | P03736 bacterioph  |
| 49 | 37   | 38.5 | 875  | 1 | CND3_SCHPO | 010429 schizosacch |
| 50 | 37   | 38.5 | 1082 | 1 | SP23_YEAST | P35210 saccharomyc |

RP SEQUENCE FROM N.A.  
RX MEDLINE=90253615; PubMed=1692709;  
RA Sawada R., Ohashi K., Anaguchi H., Okazaki H., Hattori M., Minato N.,  
RT Naruto M.;  
RP "Isolation and expression of the full-length cDNA encoding CD59  
RL antigen of human lymphocytes.";  
RN DNA Cell Biol. 9:213-220(1990).  
[6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92390353; PubMed=1381503;  
RA Petranka J.G., Fleener D.E., Sykes K., Kaufman R.E., Rosse W.F.;  
RT "Structure of the CD59-encoding gene: further evidence of a  
RL relationship to murine lymphocyte antigen Ly-6 protein.";  
RN Proc. Natl. Acad. Sci. U.S.A. 89:7876-7879(1992).  
[7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Blood;  
RX MEDLINE=93021133; PubMed=1383553;  
RA Tone M., Walsh U.A., Waldmann H.;  
RT "Gene structure of human CD59 and demonstration that discrete mRNAs  
RL are generated by alternative polyadenylation.";  
RN J. Mol. Biol. 227:971-976(1992).  
[8]  
RP SEQUENCE OF 27-128 FROM N.A.  
RX MEDLINE=89386002; PubMed=2476718;  
RA Sawada R., Ohashi K., Okano K., Hattori M., Minato N., Naruto M.;  
RT "Complementary DNA sequence and deduced peptide sequence for  
RL CD59/MEM-43 antigen, the human homologue of murine lymphocyte antigen  
Ly-6C.";  
RN Nucleic Acids Res. 17:6728-6728(1989).  
[9]  
RP GPI-ANCHOR AND DISULFIDE BONDS.  
RX MEDLINE=94103166; PubMed=8276756;  
RA Sugita Y., Nakano Y., Oda E., Noda K., Tobe T., Miura N.H., Tomita M.;  
RT "Determination of carboxyl-terminal residue and disulfide bonds of  
RL MAC1F (CD59), a glycosyl-phosphatidylinositol-anchored membrane  
protein.";  
RN J. Biochem. 114:473-477(1993).  
[10]  
RP BINDING STUDIES.  
RX MEDLINE=92317101; PubMed=1377690;  
RA Nimolaya H., Sims P.J.;  
RT "The human complement regulatory protein CD59 binds to the alpha-  
RL chain of C8 and to the 'b' domain of C9.";  
RN J. Biol. Chem. 267:13675-13680(1992).  
[11]  
RP MUTATIONAL ANALYSIS.  
RX MEDLINE=97188784; PubMed=9053451;  
RA Bodian D.L., Davis S.J., Morgan B.P., Rushmere N.K.;  
RT "Mutational analysis of the active site and antibody epitopes of the  
RL complement-inhibitory glycoprotein, CD59.";  
RN J. Exp. Med. 185:507-516(1997).  
[12]  
RP IDENTIFICATION OF COMPLEMENT INHIBITORY DOMAIN.  
RX MEDLINE=97383147; PubMed=9235986;  
RA Yu J., Dong S., Rushmere N.K., Morgan B.P., Abegyan R., Tomlinson S.;  
RT "Mapping the regions of the complement inhibitor CD59 responsible for  
RL its species selective activity.";  
RN Biochemistry 36:9423-9428(1997).  
[13]  
RP STRUCTURE OF CARBOHYDRATES AND GPI-ANCHOR, AND SEQUENCE OF N-TERMINUS.  
RX MEDLINE=97207284; PubMed=9054419;  
RA Rudd P.M., Morgan B.P., Wormald M.R., Harvey D.J., van den Berg C.W.,  
RT "The glycosylation of the complement regulatory protein, human  
RL erythrocyte CD59.";  
RN J. Biol. Chem. 272:7229-7244(1997).  
[14]  
RP INHIBITION BY GLYCATION, AND MUTAGENESIS OF LYS-66 AND HIS-69.  
RX MEDLINE=20266386; PubMed=10805801;  
RA Acosta J., Hettler J., Flueckiger R., Krumrei N., Goldfine A.,  
RT "Molecular basis for a link between complement and the vascular  
RN complications of diabetes.";  
RN Proc. Natl. Acad. Sci. U.S.A. 97:5450-5455(2000).  
[15]  
RP STRUCTURE BY NMR OF 26-95.  
RX MEDLINE=94213818; PubMed=7512825;  
RA Kiefer B., Driscoll P.C., Campbell I.D., Willis A.C.,  
RT "Three-dimensional solution structure of the extracellular region of  
RL the complement regulatory protein CD59, a new cell-surface protein  
RN domain related to snake venom neurotoxins.";  
RN Biochemistry 33:4471-4482(1994).  
[16]  
RP STRUCTURE BY NMR OF 26-102.  
RC TISSUE=Urine;  
RX MEDLINE=94348877; PubMed=7520819;  
RA Fletcher C.M., Harrison R.A., Lachmann P.J., Neuhaus D.;  
RT "Structure of a soluble, glycosylated form of the human complement  
RL regulatory protein CD59.";  
RN Structure 2:185-199(1994).  
- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C8 AND/OR C9  
INCORPORATION OF THE ASSEMBLING MAC, THEREBY PREVENTING  
FORMATION OF THE OSMOTIC PORE. THIS INHIBITOR APPEARS TO BE  
SPECIES-SPECIFIC. INVOLVED IN SIGNAL TRANSDUCTION FOR T-CELL  
ACTIVATION COMPLEXED TO A PROTEIN TYROSINE KINASE. INTERACTS WITH  
T-CELL SURFACE ANTIGEN CD2.  
- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
- PGM: N-AND O-GLYCOSYLATED. THE N-GLYCOSYLATION MAINLY CONSISTS OF A  
FAMILY OF BI-ANTENNARY COMPLEX-TYPE STRUCTURES WITH AND WITHOUT  
LACTOSAMINE EXTENSIONS AND OUTER ARM FUCCOSE RESIDUES. THE  
PREDOMINANT O-GLYCANS ARE MONO-SIALYLATED FORMS OF THE  
DISACCHARIDE, GAL-BETA-1,3GALNAc, AND THEIR SITES OF ATTACHMENT  
ARE PROBABLY ON THR-76 AND THR-77.  
- PGM: GLYCATED. GLYCATION IS FOUND IN DIABETIC SUBJECTS, BUT ONLY  
AT MINIMAL LEVELS IN NONDIABETIC SUBJECTS. GLYCATED CD59 LACKS  
MAC-INHIBITORY FUNCTION AND CONFERS TO VASCULAR COMPLICATIONS OF  
DIABETES.  
- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
- DATABASE: NAME=PRO; NOTE=CD guide CD59 entry;  
WWW="http://www.ncbi.nlm.nih.gov/ncbi/prow/cd/cd59.htm".  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
DR EMBL; M27909; AAA60543.1; -  
DR EMBL; M95708; AAA60957.1; -  
DR EMBL; X16447; CAA34467.1; -  
DR EMBL; X17198; CAA35059.1; -  
DR EMBL; X15861; CAA33870.1; -  
DR EMBL; M34671; AAA51952.1; -  
DR EMBL; M84345; - NOT\_ANNOTATED\_CDS.  
DR EMBL; M84345; - NOT\_ANNOTATED\_CDS.  
DR EMBL; M84346; AAA88793.1; JOINED.  
DR EMBL; M84348; AAA88793.1; JOINED.  
DR EMBL; Z14113; CAA78486.1; -  
DR EMBL; Z14114; CAA78486.1; -  
DR EMBL; Z14115; CAA78486.1; JOINED.  
DR PIR; J10109; RWH059.  
DR PIR; A34587; A34587.  
DR PIR; S05504; S05504.  
DR PIR; S09201; S09201.  
DR PIR; A46252; A46252.  
DR PDB; 1CDQ; 30-SEP-94.  
DR PDB; 1CDR; 30-SEP-94.  
DR PDB; 1CDG; 30-SEP-94.  
DR PDB; 1ERG; 30-APR-94.  
DR PDB; 1ERH; 30-APR-94.

DR GLycoSuiteDB; P13987; -  
 DR MIM; 107271; -  
 DR InterPro: IPR001526; LY6\_UPAR.  
 DR Pfam: PF00021; UPAR\_LY6; 1.  
 DR SMART; SM00134; LU; 1.

Query Match 100.0%; Score 96; DB 1; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFNDVTTLRRE 17  
 Db 67 FEHCNFNDVTTLRRE 83

## RESULT 2

CD59\_PAPSP STANDARD; PRT; 126 AA.  
 AC Q28785;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION  
 DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).  
 GN CD59.  
 OS Papio sp. (Baboon).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Papio.  
 NCBI\_TaxID=61183;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95104908; PubMed=7528724;  
 RA Fodor W.L., Rollins S.A., Bianco-Caron S., Burton W.V.,  
 RA Guilmette E.R., Rother R.P., Zavolico G.B., Squitto S.P.;  
 RT "Primate terminal complement inhibitor homologues of human CD59.";  
 RL Immunogenetics 41:51-51(1995).  
 CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
 CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC  
 CC ASSEMBLY.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L22862; AAA74127.1; -  
 DR HSSP; P13987; ICDO.  
 DR InterPro: IPR001526; LY6\_UPAR.  
 DR Pfam: PF00021; UPAR\_LY6; 1.  
 DR SMART; SM00134; LU; 1.  
 DR PROSITE; PS00983; LY6\_UPAR; 1.  
 DR Antigen; Glycoprotein; GPI-anchor; Signal.  
 KW Antigen; Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 1 100  
 FT PROPEP 101 126  
 FT DOMAIN 26 106  
 FT DISULFID 28 106  
 FT DISULFID 31 38  
 FT DISULFID 44 64  
 FT DISULFID 70 88  
 FT DISULFID 89 94  
 FT CARBOHYD 43 43  
 FT LIPID 100 100  
 FT SEQUENCE 126 AA; 13716 MW; 7900FF937871EBDC CRC64;

Query Match 65.6%; Score 63; DB 1; Length 126;  
 Best Local Similarity 58.8%; Pred. No. 0.00095;  
 Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNFNDVTTLRRE 17  
 Db 67 FANCNFNDISTLLKESE 83

## RESULT 3

CD59\_CERAE STANDARD; PRT; 128 AA.  
 AC Q28216;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION  
 DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).  
 GN CD59.  
 OS Cercopithecus aethiops (Green monkey) (Givet).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 NCBI\_TaxID=9534;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95104908; PubMed=7528724;  
 RA Fodor W.L., Rollins S.A., Bianco-Caron S., Burton W.V.,  
 RA Guilmette E.R., Rother R.P., Zavolico G.B., Squitto S.P.;  
 RT "Primate terminal complement inhibitor homologues of human CD59.";  
 RL Immunogenetics 41:51-51(1995).  
 CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
 CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC  
 CC ASSEMBLY.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L22863; AAA74126.1; -  
 DR HSSP; P13987; ICDO.  
 DR InterPro: IPR001526; LY6\_UPAR.  
 DR Pfam: PF00021; UPAR\_LY6; 1.  
 DR SMART; SM00134; LU; 1.  
 DR PROSITE; PS00983; LY6\_UPAR; 1.  
 DR Antigen; Glycoprotein; GPI-anchor; Signal.  
 KW Antigen; Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 102  
 FT PROPEP 103 128  
 FT LIPID 102 102  
 FT DOMAIN 26 108  
 FT DISULFID 28 51  
 FT DISULFID 31 38  
 FT DISULFID 44 64  
 FT DISULFID 70 88  
 FT DISULFID 89 94  
 FT CARBOHYD 43 43  
 FT SEQUENCE 128 AA; 14007 MW; 9778DEF7F705152 CRC64;

Query Match 65.6%; Score 63; DB 1; Length 128;  
 Best Local Similarity 58.8%; Pred. No. 0.00096;  
 Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNFNDVTTLRRE 17  
 Db 67 FANCNFNDISTLLKESE 83

Db 67 FANCNENDISTLKRESE 83

```
RESULT 4
CD59_RABIT
ID CD59_RABIT STANDARD: PRT: 124 AA.
AC 077541;
DT 20-AUG-2001 (Rel. 40, Created)
DE 20-AUG-2001 (Rel. 40, last sequence update)
DE CD59_GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN NCBI_TaxId=9986;
RP SEQUENCE FROM N.A. AND SEQUENCE OF 25-64.
RX MEDLINE=98221206; PubMed=955129;
RA Zhao X.-J., Zhou J., Zhou Q., Sims P.J.;
RT "Identity of the residues responsible for the species-restricted
RT complement inhibitory function of human CD59."
CC J. Biol. Chem. 273:10665-10671(1998).
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
CC COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C8 AND/OR C9
CC INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR COMPLETE
CC FORMATION OF THE OSMOTIC PORE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (BY SIMILARITY).
CC -1- MISCELLANEOUS: THE MATURE FORM OF THIS CD59 CONTAINS AN ADDITIONAL
CC SERINE RESIDUE BEFORE THE CONSERVED N-TERMINAL LEUCINE RESIDUE
CC FOUND IN ALL OTHER CD59 HOMOLOGS SEQUENCED TO DATE.
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF040387; AAC23590.1; -
DR HSSP: P13987; 1ERG.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LY; 1.
DR PROSITE: PS00983; LY6_UPAR; FALSE_NEG.
KW Antigen; Glycoprotein; GPI-anchor; signal.
FT SIGNAL 1 24
FT CHAIN 25 101
FT PROPEP 102 124 CD59_GLYCOPROTEIN.
FT DOMAIN 25 101 REMOVED IN MATURE FORM (BY SIMILARITY).
FT DISULFID 28 51 UPAR/LY6.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 44 64 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 89 94 BY SIMILARITY.
FT CARBOHYD 37 37 N-LINKED (GLCNAC... ) (POTENTIAL).
FT LIPID 101 101 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 124 AA; 13870 MW; CEA64C81677ZCABD CRC64;

Query Match
Best Local Similarity 51.0%; Score 49; DB 1; Length 124;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FHCNENFNDYTRRLRENE 16
DB 67 YEDCNFERSINRLSEN 82
```

```
RESULT 5
CD59_AOTR
ID CD59_AOTR STANDARD: PRT: 128 AA.
AC P51447;
DT 01-OCT-1996 (Rel. 34, Created)
DE 01-OCT-1996 (Rel. 34, last sequence update)
DE 01-NOV-1997 (Rel. 35, last annotation update)
DE CD59_GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
OS Aotus trivirgatus (Night monkey) (Douroucouli).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
RN NCBI_TaxId=9505;
RP SEQUENCE FROM N.A.
RX MEDLINE=95104908; PubMed=7528724;
RA Fodor W.L., Rollins S.A., Blanco-Caron S., Burton W.V.,
RA Guilmette E.R., Rother R.P., Zavolco G.B., Squinto S.P.;
RT "Primate terminal complement inhibitor homologues of human CD59."
RT Immunogenetics 41:51-51(1995).
CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5B-8 STAGE OF MAC
CC ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L22861; AAA3372.1; -
DR HSSP: P13987; 1CDO.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LY; 1.
DR PROSITE: PS00983; LY6_UPAR; 1.
KW Antigen; Glycoprotein; GPI-anchor; signal.
FT SIGNAL 1 25
FT CHAIN 26 102
FT PROPEP 103 128 CD59_GLYCOPROTEIN.
FT DOMAIN 26 108 REMOVED IN MATURE FORM (BY SIMILARITY).
FT DISULFID 28 51 UPAR/LY6.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 44 64 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 89 94 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC... ) (POTENTIAL).
FT LIPID 102 102 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 128 AA; 14200 MW; 62D219B5589E55B CRC64;

Query Match
Best Local Similarity 51.0%; Score 49; DB 1; Length 128;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FHCNENFNDYTRRLRENE 17
DB 67 FEDCTFSRVSQNSLENE 83

RESULT 6
CD59_CALSO
ID CD59_CALSO STANDARD: PRT: 128 AA.
AC P46557;
DT 01-NOV-1995 (Rel. 32, Created)
```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE COMPLEX INHIBITION
DE FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
GN CD59.
OS Callitrix sp. (Marmoset).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrix.
OX NCBI_TaxId=9485;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95104908; PubMed=7528724;
RA Fodor W.L., Rollins S.A., Bianco-Caron S., Burton W.V.,
RA Guilmette E.R., Rother R.P., Zavoico G.B., Squitto S.P.;
RT "Primate terminal complement inhibitor homologues of human CD59.";
CC Immunogenetics 41:51-51(1995).
CC -I- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
CC COMPLEX (MAC) ACTION. ACTS AT OR AFTER THE C5b-8 STAGE OF MAC
CC ASSEMBLY.
CC -I- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
CC (BY SIMILARITY).
CC -I- SIMILARITY: CONAINS 1 UPAR/LY6 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L2860; AAA35398.1; .
DR HSSP: P13987; ICDO.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LU; 1.
DR PROSITE: PS00983; LY6_UPAR; 1.
KW Antigen; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 25
FT CHAIN 26 102 CD59 GLYCOPROTEIN.
FT PROEP 103 128 REMOVED IN MATURE FORM (BY SIMILARITY).
FT DOMAIN 26 108 UPAR/LY6.
FT DISULFID 28 51 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 44 64 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 89 94 BY SIMILARITY.
FT CARBOHYD 43 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 102 102 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 128 AA; 14210 MW; 7A44CCAADECEDB4 CRC64;

Query Match 49.0%; Score 47; DB 1; Length 128;
Best Local Similarity 47.1%; Pred. No. 0.47;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FHCNPNVDYTRLRENE 17
DB 67 FEDCTFRLSNLSENE 83

RESULT 7
YVS5 CAEEL
ID YVS5 CAEEL STANDARD; PRT; 464 AA.
AC Q10011;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOHETICAL 51.9 KDA PROTEIN T19C3.5 IN CHROMOSOME III PRECURSOR.
GN T19C3.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;

```

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Fulton L.;
RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: SOME, TO THE BPI/CEP/LBP/PLTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28412; AAC46596.1; .
DR WormPep: T19C3.5; CE02036.
DR InterPro: IPR001124; LBP_BPI_CETP.
DR Pfam: PF01273; LBP_BPI_CETP; 1.
DR SMART: SM00328; BPI1; 1.
DR SMART: SM00329; BPI2; 1.
DR PROSITE: PS00400; LBP_BPI_CETP; FALSE_NEG.
KW Hypothetical protein; Signal.
FT CHAIN 1 17
FT SIGNAL 18 464 POTENTIAL.
FT SIGNAL 18 464 HYPOTHETICAL PROTEIN T19C3.5.
SQ SEQUENCE 464 AA; 51939 MW; 3029CFD2F65E67F CRC64;

Query Match 43.8%; Score 42; DB 1; Length 464;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 HCNPNVDYTRLREN 16
DB 108 HAKETVYTRLKRN 121

RESULT 8
ATPF_ODOSI
ID ATPF_ODOSI STANDARD; PRT; 179 AA.
AC Q00822;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE B CHAIN (EC 3.6.1.34) (SUBUNIT 1).
GN ATPF.
OS Odontella sinensis.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Eudiplomphycidae; Eupodiscaceae; Odontella.
OX NCBI_TaxId=2839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219274; PubMed=1532839;
RA Pantic P.G., Strotmann H., Kowallik K.V.;
RT "Chloroplast ATPase genes in the diatom Odontella sinensis reflect
RT cyanobacterial characters in structure and arrangement.";
RN J. Mol. Biol. 224:529-536(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pantic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RN Plant Mol. Biol. Rep. 13:336-342(1995).
RN [3]
RP SEQUENCE OF 141-179 FROM N.A.
RX MEDLINE=91192176; PubMed=1826484;
RA Pantic P.G., Strotmann H., Kowallik K.V.;
RT "The delta subunit of the chloroplast ATPase is plastid-encoded in
RT the diatom Odontella sinensis.";
RN FEBS Lett. 280:387-392(1991).

```

CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), Epsilon(1). CF(0)  
 CC HAS THREE MAIN SUBUNITS: A, B AND C.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X60752; CAA43155.1; -  
 CC EMBL: 267753; CAA91692.1; -  
 CC EMBL: X57701; -; NOT\_ANNOTATED\_CDS.  
 CC PIR: S14444; S14444.  
 CC PIR: S23358; S23358.  
 CC Mendel: 4844; ODOsi:atpf.1  
 CC InterPro: IPR002146; ATP-synt\_B.  
 CC Pfam: PF00430; ATP-synt\_B.1.  
 CC KJ: Transmembrane protein, 24  
 CC TRANSMEM 24  
 CC PT Hydrogen ion transport, 42  
 CC SQ SEQUENCE 179 AA; 20074 MW; 4B614F93D501483D CRC64; POTENTIAL.

Query Match  
 Best Local Similarity 42.7%; Score 41; DB 1; Length 179;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 5 NFNVDVTRLENE 17  
 Db 3 NFNQIFTLAENE 15  
 ||| : ||| : ||| :  
 ||| : ||| : ||| :

RESULT 9  
 COAT\_TPRCA  
 ID COAT\_TPRCA STANDARD: PRT; 223 AA.  
 AC P05070;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DE COAT PROTEIN (CAPSID PROTEIN).  
 OS Tobacco rattle virus (strain CAM).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobravirus.  
 OX NCBI\_TaxID=12296;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86093657; PubMed=3841203;  
 RA Bergin S.T., Koziel M.G., Huang S.-C., Thomas R.A., Gilley D.P.,  
 RA Siegel A.;  
 RT "The nucleotide sequence of tobacco rattle virus RNA-2 (CAM strain).";  
 RL Nucleic Acids Res. 13:8507-8518(1985).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X03241; CAA26998.1; -  
 CC EMBL: A26027; VCBVCA.  
 CC PIR: A26027; VCBVCA.  
 CC DR Coat protein.  
 CC KW Coat protein.  
 CC SQ SEQUENCE 223 AA; 23683 MW; EF1652A2F743398 CRC64;

Query Match  
 Best Local Similarity 42.7%; Score 41; DB 1; Length 223;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 NFNVDVTRLR 14  
 Db 25 NFKVDVTRLR 34  
 ||| : ||| : ||| :  
 ||| : ||| : ||| :

RESULT 10  
 ID YD71\_SCHPO STANDARD: PRT; 720 AA.  
 AC Q10326;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE HYPOHETICAL 82.5 KDA PROTEIN C32A11.01 IN CHROMOSOME I.  
 GN SPAC32A11.01.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX SKELTON J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.,  
 RL Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.,  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z69796; CAA93698.1; -  
 CC KW Hypothetical protein.  
 CC SQ SEQUENCE 720 AA; 82456 MW; D4DD9B6FC6E5604C CRC64;

Query Match  
 Best Local Similarity 42.7%; Score 41; DB 1; Length 720;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FENCFNVDVTT 11  
 Db 383 FSHCFSPESLT 393  
 ||| : ||| : ||| :  
 ||| : ||| : ||| :

RESULT 11  
 ID VGL2\_IBVB STANDARD: PRT; 1162 AA.  
 AC P11223; F05134;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)  
 GN [CONAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].  
 OS Avian infectious bronchitis virus (strain Beaudette) (IBV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirinae;  
 OX NCBI\_TaxID=11122;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85159540; PubMed=2984314;  
 RA Binn M.M., Boursnell M.E.G., Cavanagh D., Pappin D.J.C.,  
 RA Brown T.D.K.;  
 RT "Cloning and sequencing of the gene encoding the spike protein of the  
 RT coronavirus IBV.";  
 RL J. Gen. Virol. 66:719-726(1985).  
 CC [2]  
 CC SQ SEQUENCE FROM N.A.  
 CC MEDLINE=87085499; PubMed=3025348;



RA Bins M.M., Boursnell M.E.G., Tomley F.M., Brown D.K.;  
 RT "Comparison of the spike precursor sequences of coronavirus IBV  
 RL strains M41 and 6/82 with that of IBV Beaudette.";  
 CC J. Gen. Virol. 67:2825-2831(1986).  
 CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS  
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: M95169; AAA70235.1; -  
 CC DR EMBL: X02342; CA26201.1; -  
 CC DR InterPro: IPR002551; Corona\_S1.  
 CC DR InterPro: IPR002552; Corona\_S2.  
 CC DR Pfam: PF01600; Corona\_S1; 1.  
 CC DR Pfam: PF01601; Corona\_S2; 1.  
 CC KW Glycoprotein; Envelope protein; Transmembrane; Signal.  
 CC FT SIGNAL 1 18  
 CC FT CHAIN 19 1162 E2 GLYCOPROTEIN.  
 CC FT CHAIN 19 537 SPIKE PROTEIN S1.  
 CC FT CHAIN 538 1162 SPIKE PROTEIN S2.  
 CC FT DOMAIN 1120 1137 CYS-RICH.  
 CC FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SQ SEQUENCE 1162 AA; 128046 MW; 0BAAD58113CEBBD5 CRC64;

Query Match Score 41; DB 1; Length 1162;  
 Best Local Similarity 75.0%; Pred. No. 50;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
 DB 101 HCNFSDTT 108

RESULT 12  
 VGL2.IBVM STANDARD; PRT; 1162 AA.  
 AC P12651;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLIMER PROTEIN)  
 DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].  
 GN S.  
 OS Avian infectious bronchitis virus (strain M41) (IBV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronaviruses.  
 OC NCBI\_TaxID=111127;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87021475; PubMed=2429473;  
 RA Niesters H.G.M., tenstra J.A., Spaan W.J.M., Zijlerveld A.J.,  
 RA Blesimink-Pluym N.M.C., Hong F., van Schatzenburg G.J.M.,  
 RA Horzinek M.C., van der Zeijst B.A.M.;  
 RT "The peplimer protein sequence of the M41 strain of coronavirus IBV  
 RT and its comparison with Beaudette strains.";  
 RT Virus Res. 5:253-263(1986).  
 CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS  
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: M21883; AAA66575.1; -  
 CC DR EMBL: A24863; CA01736.1; -  
 CC DR PIR: S07421; S07421.  
 CC DR InterPro: IPR002551; Corona\_S1.  
 CC DR InterPro: IPR002552; Corona\_S2.  
 CC DR Pfam: PF01600; Corona\_S1; 1.  
 CC DR Pfam: PF01601; Corona\_S2; 1.  
 CC KW Glycoprotein; Envelope protein; Transmembrane; Signal.  
 CC FT SIGNAL 1 18  
 CC FT CHAIN 19 1162 E2 GLYCOPROTEIN.  
 CC FT CHAIN 19 537 SPIKE PROTEIN S1.  
 CC FT CHAIN 538 1162 SPIKE PROTEIN S2.  
 CC FT DOMAIN 1120 1137 CYS-RICH.  
 CC FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SQ SEQUENCE 1162 AA; 128077 MW; 3C9CC70938492DDA CRC64;



-1- SIMILARITY: STRONG TO S.POMBE EXO2; SOME, TO YEAST RAF1 AND TO S.POMBE DHPI.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M90097; AAA35219.1; -  
DR EMBL; X54717; CAA38520.1; -  
DR EMBL; M58367; AAA55036.1; -  
DR EMBL; M36725; AAA5125.1; -  
DR EMBL; X61181; CAA43487.1; -  
DR EMBL; X84705; CAA59180.1; -  
DR EMBL; Z72695; CAA96885.1; -  
DR PIR; S13743; S13743.  
DR PIR; S16885; S16885.  
DR PIR; S16701; S16701.  
DR PIR; A39790; A39790.  
DR SGP; S0003141; KEM1.  
KW Hydrolyase; Nuclease; Exonuclease; Magnesium; DNA recombination;  
KW DNA-binding; DNA damage.  
SQ SEQUENCE 1528 AA; 175459 MW; 49C2EDAF73D3EB92 CRC64;  
  
Query Match 42.7%; Score 41; DB 1; Length 1528;  
Best Local Similarity 60.0%; Pred. No. 67;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
OY 3 HCNFNDVTRLENE 17  
| : ||| ||| |  
DB 45 HCNDDVTKRLTEEE 59  
  
RESULT 14  
RPO\_CGMVS STANDARD: PRT; 1648 AA.  
ID RPO\_CGMVS STANDARD: PRT; 1648 AA.  
AC P19523; O83208; P89877; P90356;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (186 KDA PROTEIN) [CONTAINS:  
DE METHYLTRANSFERASE/RNA HELICASE (MT/HEL) (129 KDA PROTEIN)].  
OS Cucumbe green mottle mosaic virus (watermelon strain SH) (CGMV), and  
OS Cucumbe green mottle mosaic virus (watermelon strain W) (CGMW).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
OX NCBI\_TaxID=12236, 12237;  
RN NCB1  
RP SEQUENCE FROM N.A.  
RC STRAIN-SH;  
RX MEDLINE=91311400; PubMed=1856687;  
RA Ugaki M., Tomiyama M., Kakutani T., Hidaka S., Kiguchi T., Nagata R.,  
RA Sato T., Motoyoshi F., Nishiguchi M.;  
RT "The complete nucleotide sequence of cucumber green mottle mosaic  
RT virus (SH strain) genomic RNA."  
RL J. Gen. Virol. 72:1487-1495(1991).  
RN [2]  
RP REVISIONS.  
RC STRAIN-SH;  
RA Ugaki M., Tomiyama M., Kakutani T., Hidaka S., Kiguchi T., Nagata R.,  
RA Sato T., Motoyoshi F., Nishiguchi M.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1496-1648 FROM N.A.  
RC STRAIN-W;  
RX MEDLINE=89073773; PubMed=3201760;  
RA Sato T., Imai Y., Meshi T., Okada Y.;  
RT "Intertrial homologues of the 30k proteins of tobamoviruses."  
RL Virology 167:653-656(1988).  
CC -1- FUNCTION: THE LONGER PROTEIN IS A RNA POLYMERASE ACTIVE IN VIRAL

RNA REPLICATION.  
-1- FUNCTION: THE SMALLER PROTEIN IS A METHYLTRANSFERASE ACTIVE IN RNA  
CAPING AND AN RNA HELICASE.  
-1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN  
COODONS FOR LYS-1144 AND GLN-1146.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; D12505; BAA18895.1; -  
DR EMBL; D12505; BAA18896.1; -  
DR EMBL; J04322; AAA46382.1; -  
DR PIR; JQ1157; WMTMS2.  
DR InterPro; IPR001788; RNA\_dep\_RNAPol2.  
DR InterPro; IPR002588; V\_methyltransf.  
DR InterPro; IPR006006; Viral\_helcsl.  
DR Pfam; PF00978; RNA\_dep\_RNAPol2; 1.  
DR Pfam; PF01443; Viral\_helcsl; 1.  
DR Pfam; PF01660; Vmethyltransf; 1.  
KW Transferase; RNA-directed RNA polymerase; Helicase; ATP-binding.  
FT CHAIN 1 1648  
FT CHAIN 1 1144  
FT NP\_BIND 863 870  
FT NP\_BIND 1 1144  
SQ SEQUENCE 1648 AA; 186549 MW; 1D1AFEEEB7B65595 CRC64;  
  
Query Match 42.7%; Score 41; DB 1; Length 1648;  
Best Local Similarity 61.5%; Pred. No. 72;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
OY 5 NENDVTRLENE 17  
| : ||| ||| |  
DB 1188 DFDVATMRLRDNE 1200  
  
RESULT 15  
UTRO\_HUMAN STANDARD: PRT; 3433 AA.  
ID UTRO\_HUMAN STANDARD: PRT; 3433 AA.  
AC P46939;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE UTRORPHIN (DYSTROPHIN-RELATED PROTEIN 1) (DRP1) (DRP).  
GN UTRN OR DMDL.  
OS Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93096045; PubMed=1461283;  
RA Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J.,  
RA Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R.,  
RA Edwards Y.H., Davies K.E.;  
RT "Primary structure of dystrophin-related protein."  
RL Nature 360:591-593(1992).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.  
RX MEDLINE=99141377; PubMed=9887274;  
RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J.,  
RA Kendrick-Jones J.;  
RT "The 2.0-A structure of the second calponin homology domain from the  
RT actin-binding region of the dystrophin homologue utrophin."  
RL J. Mol. Biol. 285:1257-1264(1999).  
CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE  
CC PLASMA MEMBRANE (BY SIMILARITY TO DYSTROPHIN).  
CC -1- SUBCELLULAR LOCATION: NEURONMUSCULAR JUNCTION.  
CC -1- TISSUE SPECIFICITY: MUSCLE.

CC -1- SIMILARITY: STRONG, TO DYSTROPHIN.  
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY  
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,  
 CC ABP-120, ABP-180, OR BETA-FODRIN).  
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 20 SPECTRIN REPEATS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X69086; CAA48829.1;  
 CC PIR: S24381; S28381.  
 CC PDB: 1BHD; 16-FEB-99.  
 CC MTM: 128240;  
 CC InterPro: IPR001589; Actinin\_act\_bind.  
 CC InterPro: IPR001715; Calponin\_hom.  
 CC InterPro: IPR003122; Spectrin.  
 CC InterPro: IPR001202; WW.  
 CC InterPro: IPR003349; WW\_Domain.  
 CC Pfam: PF00307; CH; 2.  
 CC Pfam: PF00435; Spectrin; 19.  
 CC Pfam: PF00397; WW; 1.  
 CC Pfam: PF00369; Z2; 1.  
 CC PRINTS: PR00403; WMDOMAIN.  
 CC SMART: SM00033; CH; 2.  
 CC SMART: SM00150; SPEC; 18.  
 CC SMART: SM00319; Tarr; 1.  
 CC SMART: SM00456; WW; 1.  
 CC PROSITE: PS00019; ACTININ\_1; 1.  
 CC PROSITE: PS00020; ACTININ\_2; 1.  
 CC PROSITE: PS01159; CH; 2.  
 CC PROSITE: PS00020; WW\_DOMAIN\_2; 1.  
 CC DR Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;  
 CC Repeat; 3D-structure.  
 CC FT DOMAIN 1 246  
 CC FT REPEAT 246 364  
 CC FT REPEAT 309 417  
 CC FT REPEAT 418 526  
 CC FT REPEAT 541 637  
 CC FT REPEAT 687 798  
 CC FT REPEAT 803 902  
 CC FT REPEAT 1016 1083  
 CC FT REPEAT 1125 1230  
 CC FT REPEAT 1248 1334  
 CC FT REPEAT 1432 1541  
 CC FT REPEAT 1544 1649  
 CC FT REPEAT 1652 1753  
 CC FT REPEAT 1910 1968  
 CC FT REPEAT 1976 2081  
 CC FT REPEAT 2258 2333  
 CC FT REPEAT 2399 2440  
 CC FT REPEAT 2443 2556  
 CC FT REPEAT 2559 2636  
 CC FT REPEAT 2658 2688  
 CC FT REPEAT 2691 2797  
 CC FT DOMAIN 2812 2845  
 CC FT DOMAIN 3064 3111  
 CC FT DOMAIN 3064 3117  
 CC SEQUENCE 3433 AA; 394488 MW; EA8BD409F8585B CRC64;  
 CC CYS-RICH.

Query Match  
 Best Local Similarity 42.7%; Score 41; DB 1; Length 3433;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 HCNENDVYTRLENE 17  
 Db 1982 HCDLNDLTOWITEAE 1996

RESULT 16  
 ID R1M4\_RICPR STANDARD; PRT; 165 AA.  
 AC Q9ZDI0;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE PROBABLE 16S RRNA PROCESSING PROTEIN R1M4.  
 GN R1M4 OR RP348.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 NC NCB1\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MADRID E;  
 RA MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Sierhartz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kuiland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RU mitochondria.";  
 RL Nature 396:133-140(1998).  
 CC -1- FUNCTION: ESSENTIAL FOR EFFICIENT PROCESSING OF 16S RRNA. PROBABLY  
 CC PART OF THE 30S SUBUNIT PRIOR TO OR DURING THE FINAL STEP IN THE  
 CC PROCESSING OF 16S FREE 30S RIBOSOMAL SUBUNITS. IT COULD BE SOME  
 CC ACCESSORY PROTEIN NEEDED FOR EFFICIENT ASSEMBLY OF THE 30S  
 CC SUBUNIT. R1M4 IS NEEDED IN A STEP PRIOR TO RRNA DURING THE  
 CC MATURATION OF 16S RRNA. HAS AFFINITY FOR FREE RIBOSOMAL 30S  
 CC SUBUNITS BUT NOT FOR 70S RIBOSOMES (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE R1M4 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: AJ235271; CAA14808.1;  
 CC InterPro: IPR002676; R1M4.  
 CC DR Pfam: PF01782; R1M4; 1.  
 CC KW rRNA processing; Complete proteome.  
 CC SEQUENCE 165 AA; 18772 MW; 25DDB377BC250919 CRC64;

Query Match  
 Best Local Similarity 41.7%; Score 40; DB 1; Length 165;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CNEFDVYTR 12  
 Db 63 CTFNDIATR 71

RESULT 17  
 ID HFA3\_HAEIN STANDARD; PRT; 212 AA.  
 AC P43988;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DE MAJOR FIMBRIAL SUBUNIT PRECURSOR (PILIN).  
 GN HIFA.  
 OS Haemophilus influenzae.

```

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86-0295 / LKP SEROTYPE 1;
RA Green B.A., Olmsted S.B.;
RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MEDIATES ADHERENCE TO OROPHARYNGEAL EPITHELIAL CELLS.
CC HELPS THE AIRWAY COLONIZATION PROCESS.
CC -1- SUBCELLULAR LOCATION: FIMBRIA.
CC -1- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19730; AAA61814.1; -
DR InterPro: IPR000259; Fimbrin.
DR Pfam: PF00419; Fimbrin; 1.
KW Fimbrin; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 212 MAJOR FIMBRIAL SUBUNIT.
FT DISULFID 42 82 PROBABLE.
FT SEQUENCE 212 AA; 23003 MW; 3FABAB87BA6901B1 CRC64;

Query Match 41.7%; Score 40; DB 1; Length 212;
Best Local Similarity 47.1%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 FEHCNFDVYTRLENE 17
   | | | | | | | | | |
DB 151 FTHSNTNDVATQOTYVK 167

RESULT 18
FLPA_SULAC STANDARD; PRT; 231 AA.
AC Q9PM0;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FIBRILARIN-LIKE PRE-RRNA PROCESSING PROTEIN.
GN FLPA OR FIB.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2285;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20337948; PubMed=10775111;
RA Omer A.D., Love T.M., Russell A.G., Ebhardt H., Eddy S.R.,
RA Dennis P.P.;
RT "Homologs of small nucleolar RNAs in Archaea.";
RL Science 288:517-522(2000).
CC -1- FUNCTION: MAY BE INVOLVED IN PRE-RRNA PROCESSING.
CC -1- SIMILARITY: BELONGS TO THE FIBRILARIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF201093; AAF69254.1; -
DR InterPro: IPR000692; Fibrillar.
DR Pfam: PF01269; Fibrillar; 1.

```

```

DR PRINTS; PR00052; FIBRILARIN.
DR ProDom; PD004637; Fibrillar; 1.
DR ProSITE; PS00566; Fibrillar; 1.
KW rRNA processing; RNA-binding.
SQ SEQUENCE 231 AA; 26445 MW; 430DC184CB2EC77B CRC64;

Query Match 41.7%; Score 40; DB 1; Length 231;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CNFNDVYTRL 13
   | | | | |
DB 18 CIPNDGTRL 27

RESULT 19
KMS1_SCHPO STANDARD; PRT; 607 AA.
AC P87245;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE KARYOGAMY MEIOTIC SEGREGATION PROTEIN 1.
GN KMS1 OR SPO3A11.05C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=97294501; PubMed=9150257;
RA Shimanuki M., Miki F., Ding D.-Q., Chikashige Y., Hiraoka Y.,
RA Horio T., Niwa O.;
RT "A novel fission yeast gene, kms1+, is required for the formation of
RT meiotic prophase-specific nuclear architecture.";
RL Mol. Gen. Genet. 254:238-249(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Oliver K., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: HAS A ROLE IN KARYOGAMY, RECOMBINATION AND SEGREGATION
CC DURING MEIOSIS. ALTHOUGH IT HAS BEEN SHOWN TO ASSOCIATE WITH THE
CC SPINDLE POLE BODY IT IS UNLIKELY TO BE INVOLVED IN ITS FORMATION
CC OR MAINTENANCE SINCE KMS1(-) MUTANTS ARE ABLE TO COMPLETE MITOSIS
CC SUCCESSFULLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D84439; BAA20460.1; -
DR EMBL: Z99260; CAB16381.1; -
KW Cell division; Meiosis.
SQ SEQUENCE 607 AA; 69240 MW; 216A1D5CA93C9550 CRC64;

Query Match 41.7%; Score 40; DB 1; Length 607;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 5 NFNDVYTRLENE 17
   | | | | |
DB 70 NFRDVKYKLENE 82

RESULT 20

```

HS83\_DROAV  
ID HS83\_DROAV STANDARD: PRT: 716 AA.  
AC 002192;  
DT 15-DEC-1998 (Rel. 37, Created)  
DE 15-DEC-1998 (Rel. 37, Last sequence update)  
HEAT SHOCK PROTEIN 83 (HSP 82).  
GN HSP83.  
OS Drosophila auraria (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
RN NCBI\_TaxID=47315;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3040.11B.  
RA Konstantopoulos I., Scouras Z.G.;  
RT "The hsp83 gene of Drosophila auraria."  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY.  
CC (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: U75687; AAB58358.1;  
DR FLYBASE; FBgn0020208; Daur\Hsp83.  
DR InterPro; IPR003594; HATPase\_C.  
DR Pfam; PF02518; HATPase\_C; 1.  
DR Pfam; PF00183; HSP90; 1.  
DR PRINTS; PR00775; HEATSHOCK90.  
DR SMART; SM00387; HATPase\_C; 1.  
DR PROSITE; PS00296; HSP90; 1.  
DR Chaperone; ATP-binding; Heat shock.  
SQ SEQUENCE 716 AA; 81760 MW; 9EC59F027C4DFCD7 CRC64;  
  
Query Match  
Best Local Similarity 41.7%; Score 40; DB 1; Length 716;  
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
  
QY 2 EHCNENDVYTRRENE 17  
Db 457 DFCSLSDYVRMKENQ 472

RT transporters.";  
RL Science 257:1271-1273(1992).  
CC -1- FUNCTION: MAY ACT AS A VESICLE-LOCALIZED NEUROTRANSMITTER.  
CC TRANSPORTER OR MAY FUNCTION AS AN ION TRANSPORTER OR CHANNEL.  
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLE.  
CC -1- TISSUE SPECIFICITY: NEURAL AND ENDOCRINE CELLS OF BRAIN AND  
CC SPINAL CORD.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: L05435; AAA42188.1;  
DR Pfam; PF00083; sugar-1;  
KW Synapse; Nerve; Glycoprotein; Neurotransmitter transport;  
KW Transmembrane.  
FT DOMAIN 1  
FT TRANSMEM 164 163  
FT DOMAIN 183 182  
FT TRANSMEM 204 204  
FT TRANSMEM 225 225  
FT DOMAIN 226 242  
FT TRANSMEM 243 260  
FT DOMAIN 261 262  
FT TRANSMEM 263 281  
FT DOMAIN 282 284  
FT TRANSMEM 295 322  
FT DOMAIN 323 334  
FT TRANSMEM 335 355  
FT DOMAIN 356 445  
FT TRANSMEM 446 469  
FT TRANSMEM 470 592  
FT TRANSMEM 593 611  
FT TRANSMEM 612 626  
FT TRANSMEM 627 647  
FT TRANSMEM 648 649  
FT TRANSMEM 650 669  
FT TRANSMEM 670 694  
FT TRANSMEM 695 711  
FT TRANSMEM 712 712  
FT TRANSMEM 713 731  
FT TRANSMEM 732 742  
FT CARBOHYD 498 498  
FT CARBOHYD 548 548  
FT CARBOHYD 573 573  
SQ SEQUENCE 742 AA; 82705 MW; 565DE7EF2929D5DB CRC64;  
  
Query Match  
Best Local Similarity 41.7%; Score 40; DB 1; Length 742;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EHCNENDVYTR 11  
Db 531 FECCYFEDVTS 541

RESULT 22  
Y594\_METHA  
ID Y594\_METHA STANDARD: PRT: 169 AA.  
AC 058012;  
DT 01-NOV-1997 (Rel. 35, Created)  
DE 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
GN HYPOTHEICAL PROTEIN M0594.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
NCBI\_TaxID=2190;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RC MEDLINE-9633799; PubMed-8688087;  
 RX Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D., Overbeek A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I., Kierstead R., Kirkness E.F., Weissflog K.G., Merrick J.M., Glodek A., Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D., Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.";  
 RT Science 273:1058-1073(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U67508; AAB9595.1; -  
 CC TIGR: M0594; -  
 DR InterPro: IPR002799; IMP4.  
 DR Pfam: PF01945; IMP4; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 169 AA; 20053 MW; 9FF6090865C6E4B0 CRC64;

Query Match 41.1%; Score 39.5; DB 1; Length 169;  
 Best Local Similarity 42.1%; Pred. No. 11;  
 Matches 8; Conservative 4; Mismatches 4; Indels 3; Gaps 1;  
 Oy 1 FEHCNFN--DVTTRLREN 16  
 | | | | | | | | | |  
 Db 119 FOHLNINEDSITRLREKD 137

RESULT 23  
 ID SPOT\_SYNY3 STANDARD; PRT; 760 AA.  
 AC P74007;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROBABLE GUANOSINE-3',5'-BIS(DIPHOSPHATE) 3'-PYROPHOSPHOHYDROLASE (EC 3.1.7.2) ((PGPP)ASE) (PENTA-PHOSPHATE GUANOSINE-3'-PYROPHOSPHOHYDROLASE).  
 DE GN SPOT OR SLR1325.  
 OS Synecocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97061201; PubMed-8905231;  
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hitosawa M., Sugita M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Tabata S.;  
 RA "Sequence analysis of the genome of the unicellular cyanobacterium RT Synecocystis sp. strain PCC6803. II. Sequence determination of the RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: IN EUBACTERIA PGPP (GUANOSINE 3'-DIPHOSPHATE 5'-DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYSES THE DEGRADATION OF PGPP INTO GDP. IT MAY ALSO BE CAPABLE OF CATALYZING THE SYNTHESIS OF PGPP (BY SIMILARITY).  
 CC

CC -1- CATALYTIC ACTIVITY: GUANOSINE-3',5'-BIS(DIPHOSPHATE) + H(2)O = GUANOSINE-5'-DIPHOSPHATE + PYROPHOSPHATE.  
 CC -1- COFACTOR: MANGANESE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE RELA / SPOT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D90911; BAA18078.1; -  
 DR InterPro: IPR002912; ACT.  
 DR InterPro: IPR003607; HDC.  
 DR Pfam: PF01842; ACT; 1.  
 DR SMART: SM00471; HDC; 1.  
 KW Hydrolase; Manganese; Complete proteome.  
 SQ SEQUENCE 760 AA; 86568 MW; CA276EA2286358F7 CRC64;

Query Match 41.1%; Score 39.5; DB 1; Length 760;  
 Best Local Similarity 31.2%; Pred. No. 57;  
 Matches 10; Conservative 1; Mismatches 4; Indels 17; Gaps 1;  
 Oy 2 EHCFN-----NDVTTRLREN 16  
 | | | | | | | | | |  
 Db 548 EHCYQVVEDLACLGIGETISNSVNRREN 579

RESULT 24  
 ID ENV\_STVYZ STANDARD; PRT; 854 AA.  
 AC P17281;  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ENVELOPE POLYPROTEIN GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].  
 DE GN ENV.  
 OS Chimpanzee immunodeficiency virus (STVcpz) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90259077; PubMed-2188136;  
 RA Hueb T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;  
 RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";  
 RL Nature 345:356-359(1990).  
 CC -1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X52154; CAA36407.1; -  
 CC PIR: S09990; VCIJST.  
 DR HIV: X52154; ENVSCP2.  
 DR InterPro: IPR000328; ENV GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Polypeptide; Transmembrane; Signal.  
 KM SIGNAL. 1 30  
 FT SIGNAL 31 500 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 501 854 TRANSMEMBRANE GLYCOPROTEIN.  
 FT CHAIN 501 517 POTENTIAL.  
 FT TRANSMEM 501 517

```

FT TRANSMEM 675 693 POTENTIAL.
FT TRANSMEM 805 821 POTENTIAL.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA: 95803 MW: 2E249AFAD4F3D9B3 CRC64:

```

```

Query Match 41.1% Score 39.5; DB 1; Length 854;
Best Local Similarity 50.0%; Pred. No. 64;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

```

```

Db 153 KNCSEF-VTTLRDRE 167
2 EHCNPNVYTRTLENE 17
153 KNCSEF-VTTLRDRE 167
RESULT 25
CD59_PIG STANDARD; PRT; 123 AA.
AC 062680: 09X94; 09TR76;
DT 15-DEC-1998 (Rel. 37, Created)
DT 20-DEC-1998 (Rel. 37, Last sequence update)
DE CD59-ADG-2001 (Rel. 40, Last annotation update)
DE CD59 GLYCOPROTEIN PRECURSOR (MEMBRANE ATTACK COMPLEX INHIBITION
  FACTOR) (MACIF) (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Aortic endothelium;
RA MEDLINE-98217182; PubMed-9558099;
RT Hinchliffe S.J.; Rushmere N.K.; Hanna S.M.; Morgan B.P.;
  "Molecular cloning and functional characterization of the pig
  analogue of CD59: relevance to xenotransplantation.";
  J Immunol. 160:3924-3932(1998).
RN [2]
RP SEQUENCE OF 26-123 FROM N.A., AND FUNCTION.
RC TISSUE-Aortic endothelium;
RA MEDLINE-99023683; PubMed-9808497;
RT Maher S.E.; Pflugh D.L.; Larsen N.J.; Rothschild M.F.;
  "Structure/function characterization of porcine CD59: expression,
  chromosomal mapping, complement-inhibition, and costimulatory
  activity.";
  Transplantation 66:1094-1100(1998).
RN [3]

```

```

RP SEQUENCE OF 26-64.
RC TISSUE-Erythrocyte;
RA MEDLINE-95181826; PubMed-753195;
RT van den Berg C.W.; Harrison R.A.; Morgan B.P.;
  "A rapid method for the isolation of analogues of human CD59 by
  preparative SDS-PAGE: application to pig CD59.";
  J. Immunol. Methods 179:223-231(1995).
RN [1] FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
  COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C8 AND/OR C9
  COMPLEMENTS OF THE ASSEMBLING MAC, THEREBY PREVENTING
  INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR COMPLETE
  FORMATION OF THE OSMOTIC PORE.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES TESTED (LUNG, TESTIS
  IN LUNG AND SPLEEN, LOWEST LEVELS IN LIVER AND SKELETAL MUSCLE.
  SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation-
  use by non-profit institutions. There are no restrictions on its
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
DR EMBL: AF020302; AAC67231.1; -
DR EMBL: AF058328; AAD39837.1; -
DR InterPro: IPR001526; LY6-UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LY6; 1.
DR PROSITE: PS00983; LY6-UPAR; 1.
KW Antigen; Glycoprotein; GPI-anchor; Signal.
FT CHAIN 1 25
FT PROPEP 26 98
FT DOMAIN 26 103
FT DISULFID 28 51
FT DISULFID 31 38
FT DISULFID 44 65
FT DISULFID 71 89
FT CARBOHYD 90 95
FT CARBOHYD 43 43
FT LIPID 98 98
FT CONFLICT 26 26
FT CONFLICT 46 46
FT CONFLICT 63 63
SQ SEQUENCE 123 AA: 13790 MW: FDE2ED3F93C91321 CRC64:

```

```

Query Match 40.6% Score 39; DB 1; Length 123;
Best Local Similarity 35.3%; Pred. No. 9.9;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

```

Search completed: April 2, 2002, 09:23:23
Job time: 187 sec

```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:20:36 ; Search time 22.85 Seconds  
(without alignments)  
108.824 Million cell updates/sec

Title: US-09-020-393B-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFDVYTRLENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 50 summaries

Database :

1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 48    | 50.0        | 721    | 5     | Q27308 anopheles a |
| 2          | 46.5  | 48.4        | 1767   | 5     | Q19346 caenorhabdi |
| 3          | 46    | 47.9        | 284    | 2     | Q915M0 pseudomonas |
| 4          | 46    | 47.9        | 777    | 2     | Q9KER0 bacillus ba |
| 5          | 45    | 46.9        | 1307   | 5     | Q93554 caenorhabdi |
| 6          | 44    | 45.8        | 205    | 2     | Q9ZCG3 rickettsia  |
| 7          | 44    | 45.8        | 513    | 12    | Q82618 avian infec |
| 8          | 44    | 45.8        | 530    | 12    | O55343 avian infec |
| 9          | 44    | 45.8        | 807    | 1     | O27791 methanobact |
| 10         | 44    | 45.8        | 1101   | 5     | O9WD03 ciona intes |
| 11         | 43    | 44.8        | 544    | 12    | O82667 avian infec |
| 12         | 43    | 44.8        | 545    | 12    | O82619 avian infec |
| 13         | 43    | 44.8        | 739    | 5     | O9V9E6 dirosophila |
| 14         | 42.5  | 44.3        | 96     | 12    | O70950 human immun |
| 15         | 42.5  | 44.3        | 240    | 12    | P90081 human immun |
| 16         | 42.5  | 44.3        | 240    | 12    | P90082 human immun |
| 17         | 42.5  | 44.3        | 240    | 12    | P90085 human immun |
| 18         | 42.5  | 44.3        | 240    | 12    | P90080 human immun |
| 19         | 42.5  | 44.3        | 240    | 12    | P90083 human immun |

|    |      |      |     |    |                    |
|----|------|------|-----|----|--------------------|
| 20 | 42.5 | 44.3 | 313 | 2  | O9S091 borrelia bu |
| 21 | 42.5 | 44.3 | 679 | 12 | O9IK02 human immun |
| 22 | 42.5 | 44.3 | 684 | 12 | O9IK06 human immun |
| 23 | 42.5 | 44.3 | 852 | 12 | O73350 human immun |
| 24 | 42.5 | 44.3 | 861 | 12 | O9IK05 human immun |
| 25 | 42.5 | 44.3 | 861 | 12 | O9IK04 human immun |
| 26 | 42.5 | 44.3 | 861 | 12 | O9IK03 human immun |
| 27 | 42.5 | 44.3 | 862 | 12 | O9YK1 humulatia c  |
| 28 | 42   | 43.8 | 290 | 8  | O9G157 clostridium |
| 29 | 42   | 43.8 | 302 | 2  | O9LBM1 pictetia ac |
| 30 | 42   | 43.8 | 332 | 8  | O9TJ28 ormocarpum  |
| 31 | 42   | 43.8 | 339 | 8  | O9TJ29 ormocarpum  |
| 32 | 42   | 43.8 | 365 | 5  | O9XTF4 caenorhabdi |
| 33 | 42   | 43.8 | 395 | 8  | O9TK00 ormocarpum  |
| 34 | 42   | 43.8 | 419 | 2  | O9KG62 bacillus ba |
| 35 | 42   | 43.8 | 514 | 8  | O9TK31 ormocarpops |
| 36 | 42   | 43.8 | 514 | 8  | O9TK30 ormocarpops |
| 37 | 42   | 43.8 | 514 | 8  | O9TK29 ormocarpum  |
| 38 | 42   | 43.8 | 514 | 8  | O9TK28 ormocarpum  |
| 39 | 42   | 43.8 | 514 | 8  | O9TK27 ormocarpum  |
| 40 | 42   | 43.8 | 514 | 8  | O9TK26 ormocarpum  |
| 41 | 42   | 43.8 | 514 | 8  | O9TK25 ormocarpum  |
| 42 | 42   | 43.8 | 514 | 8  | O9TK23 dipysa flo  |
| 43 | 42   | 43.8 | 514 | 8  | O9TK19 pictetia an |
| 44 | 42   | 43.8 | 514 | 8  | O9TK01 dipysa orm  |
| 45 | 42   | 43.8 | 573 | 5  | O16166 dirosophila |
| 46 | 41.5 | 43.2 | 87  | 12 | O70912 human immun |
| 47 | 41.5 | 43.2 | 95  | 12 | O9YY30 human immun |
| 48 | 41.5 | 43.2 | 328 | 12 | O9IU90 human immun |
| 49 | 41.5 | 43.2 | 332 | 12 | O9IU01 human immun |
| 50 | 41.5 | 43.2 | 356 | 12 | P87970 human immun |

## ALIGNMENTS

| RESULT | 1  | PRELIMINARY; | PRT; | 721 AA. |
|--------|--|--------------|------|---------|
| Q27308 |  |              |      |         |
| AC     | Q27308:  |              |      |         |
| DT     | 01-NOV-1996 (TREMBLrel. 01, Created)                                 |              |      |         |
| DT     | 01-NOV-1996 (TREMBLrel. 01, Last sequence update)                    |              |      |         |
| DT     | 01-JUN-2001 (TREMBLrel. 17, Last annotation update)                  |              |      |         |
| DE     | HEAT SHOCK PROTEIN 82.   |              |      |         |
| GN     | HSP82.   |              |      |         |
| OS     | Anopheles albimanus (New world malaria mosquito).                    |              |      |         |
| OC     | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;        |              |      |         |
| OC     | Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; |              |      |         |
| OC     | Anopheles.   |              |      |         |
| OX     | NCBI_TaxID=7167;   |              |      |         |
| RN     | [1]  |              |      |         |
| RP     | SEQUENCE FROM N.A.   |              |      |         |
| RC     | STRAIN=SANTRA TECLA;   |              |      |         |
| RX     | MEDLINE=96209490; PubMed=8630537;                                    |              |      |         |
| RA     | Benedict M.O., Levine B.J., Ke Z.X., Cockburn A.F., Seawright J.A.;  |              |      |         |
| RT     | "Precise limitation of concerted evolution to ORFs in mosquito Hsp82 |              |      |         |
| RL     | genes. Mol. Biol. 5:73-79(1996).                                     |              |      |         |
| EMBL   | LA7285; AAB05638.1; -.   |              |      |         |
| DR     | EMBL; LA7285; AAB05638.1; -.   |              |      |         |
| DR     | HSP; P07900; IYER.   |              |      |         |
| DR     | InterPro: IPR003594; HATPase_C.                                      |              |      |         |
| DR     | InterPro: IPR001404; HSP90.  |              |      |         |
| DR     | Pfam: PF02518; HATPase_C; 1.   |              |      |         |
| DR     | Pfam: PF00183; HSP90; 1.   |              |      |         |
| DR     | PRINTS; PR00775; HEATSHOCK90.  |              |      |         |
| DR     | SMART; SM00387; HATPase_C; 1.  |              |      |         |
| DR     | PROSITE; PS00298; HSP90; 1.  |              |      |         |
| KW     | Heat shock.  |              |      |         |
| SO     | SEQUENCE 721 AA; 82153 MW; C71867C5610452EA CRC64;                   |              |      |         |

Query Match 50.0%; Score 48; DB 5; Length 721;

Best Local Similarity 43.8%; Pred. No. 9.1;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 EHCNPNVTTTRELRE 17  
Db 455 EYCSLNDYVGRKKNQ 470

## RESULT 2

ID 019346 PRELIMINARY; PRT; 1767 AA.  
AC 019346;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
GN F1IC1.5 PROTEIN.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Palmer S.;  
RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.  
DR EMBL: Z54270; CA91030.1; -;  
DR EMBL: Z54270; CA91031.1; -;  
DR InterPro: IPR000517; Ribosomal\_L30.  
DR PROSITE: PS00634; RIBOSOMAL\_L30.  
DR PROSITE: PS00324; VMA; 1.  
DR SMART: SM00327; VMA; 1.  
KW Alternative splicing.  
FT VARSPPLIC 1301 1302  
FT VARSPPLIC 1303 1767  
SQ SEQUENCE 1767 AA; 200015 MW; A113DCB86F5FE85 CRC64;

Query Match  
Best Local Similarity 54.5%; Score 46.5; DB 5; Length 1767;  
Matches 12; Conservative 1; Mismatches 4; Indels 5; Gaps 2;

OY 1 FEH----CN-FNDVTTTRELRE 17  
Db 601 FEHHLNTANVFNDDTITRTKENE 622

## RESULT 3

ID 0915M0 PRELIMINARY; PRT; 284 AA.  
AC 0915M0;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE PROBABLE TRANSCRIPTIONAL REGULATOR.  
GN PA0708.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Hickey C.K., Pham X.-O.T., Erwin A.L., Micosguchi S.D., Watrener P.,  
RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Feiler J., Selzer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000);

CC -1 SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.  
DR EMBL: AE004506; AAG04097.1; -;  
DR InterPro: IPR000847; HTH\_LysR.  
DR Pfam: PF00126; HTH\_1; 1.  
DR PROSITE: PS00044; HTH\_LYSR\_FAMILY; UNKNOWN.1.  
KW Complete proteome; DNA-binding; Transcription regulation.  
SQ SEQUENCE 284 AA; 31527 MW; 1P986AC7577C5C4 CRC64;

## Query Match

Best Local Similarity 47.9%; Score 46; DB 2; Length 284;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 HCFNPNVTTTRELRE 17  
Db 26 HCFVPSNITTTRELRE 40

## RESULT 4

ID 09KER0 PRELIMINARY; PRT; 777 AA.  
AC 09KER0;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE BH0789 PROTEIN.  
GN BH0789.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RT Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL: AP001509; BAB04508.1; -;  
KW Complete proteome.  
SQ SEQUENCE 777 AA; 88227 MW; 46BB843DADBF17B CRC64;

Query Match  
Best Local Similarity 47.9%; Score 46; DB 2; Length 777;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 FEHCNPNVTTTRELRE 15  
Db 136 FKHSNFDYLTQIKE 150

## RESULT 5

ID 093554 PRELIMINARY; PRT; 1307 AA.  
AC 093554; 002539.  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-JUN-1997 (TREMBlrel. 04, Last sequence update)  
DE P33A7.5 PROTEIN.  
GN P33A7.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McMurtry A.;  
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Croxall M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kristin J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 281067; CAB02976.1; -  
 SQ SEQUENCE 1307 AA; 147397 MW; EB966398D82D9B9C CRC64;

Query Match 46.9%; Score 45; DB 5; Length 1307;  
 Best Local Similarity 69.2%; Pred. No. 52;  
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 HCNFNDVTTRLREN 16  
 DB 654 CKINDTTLREN 666

RESULT 6  
 O92CG3 PRELIMINARY; PRT; 205 AA.  
 AC O92CG3;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE NADH DEHYDROGENASE 1 CHAIN J (NUOJ).  
 GN RP790.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=MADRID E;  
 RC MEDLINE=99039499; PubMed=9823893;  
 RA Anderson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria.";  
 RL Nature 396:133-140(1998).  
 DR EMBL: AJ235273; CAA15216.1; -  
 DR InterPro: IPR001457; Oxidored\_q3.  
 DR Pfam: PF00499; Oxidored\_q3; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 205 AA; 23142 MW; 035BD21245DB276A CRC64;

Query Match 45.8%; Score 44; DB 2; Length 205;  
 Best Local Similarity 66.7%; Pred. No. 12;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 HCNFNDVTTRLREN 16  
 DB 77 HCNFNDVTTRLREN 88

RESULT 7  
 O82618 PRELIMINARY; PRT; 513 AA.  
 AC O82618;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE SURFACE GLYCOPROTEIN.  
 GN SPIKE, S1.  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=CONNECTICUT;  
 RC MEDLINE=95159673; PubMed=7856318;  
 RA Wang L., Junker D., Hock L., Edlary E., Collisson E.W.;  
 RT "Evolutionary implications of genetic variations in the S1 gene of  
 RT infectious bronchitis virus.";  
 RL Virus Res. 34:327-338(1994).  
 DR EMBL: L18990; AAA74378.1; -  
 DR InterPro: IPR002551; Corona\_S1.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 SQ SEQUENCE 513 AA; 56541 MW; 2D4C1F986F6B045E CRC64;

Query Match 45.8%; Score 44; DB 12; Length 513;  
 Best Local Similarity 75.0%; Pred. No. 30;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCNFNDVT 10  
 DB 97 HCNFSDIT 104

RESULT 8  
 O55343 PRELIMINARY; PRT; 530 AA.  
 AC O55343;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE SPIKE GLYCOPROTEIN S1 SUBUNIT (FRAGMENT).  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=FLORIDA 18288;  
 RC Moore K.M., Jackwood M.W., Bennett J.D., Seal B.S.;  
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF027512; AAB87907.1; -  
 DR InterPro: IPR002551; Corona\_S1.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 FT NON\_TER 530 530  
 SQ SEQUENCE 530 AA; 58315 MW; 74CE496CE25688F5 CRC64;

Query Match 45.8%; Score 44; DB 12; Length 530;  
 Best Local Similarity 75.0%; Pred. No. 31;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCNFNDVT 10  
 DB 96 HCNFSDIT 103

RESULT 9  
 O27791 PRELIMINARY; PRT; 807 AA.  
 AC O27791;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE COLLAGENASE.  
 GN MTH1763.  
 OS Methanobacterium thermoautotrophicum.

OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA.H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 RA Aldege T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,  
 RA Spadator R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiani N., Caruso A., Bush D., Satter H., Patwell D., Prabhakar S.,  
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.,  
 RA "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RL J. Bacteriol 179:7135-7155(1997).  
 DR EMBL: AEO00931; AAB86229.1;  
 DR InterPro: IPR001539; Peptidase\_U32.  
 DR Pfam: PF01136; Peptidase\_U32.  
 DR ProDom: PD004398; Peptidase\_U32; 1.  
 DR PROSITE: PS01276; PEPTIDASE\_U32; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 807 AA; 91152 MW; 466E6FC0F7833D7 CRC64;

Query Match  
 Best Local Similarity 45.8%; Score 44; DB 1; Length 807;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 EHCNNDVTTRLE 15  
 Db 562 FRECDMNDVRSILRE 576

RESULT 10  
 ID Q9ND03 PRELIMINARY; PRT; 1101 AA.  
 AC Q9ND03;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE ATP CITRATE-LYASE.  
 GN CIT-ACL.  
 OS Clona intestinalis  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;  
 OC Clonidae; Clona.  
 OX NCBI\_TaxID=7719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,  
 RA Sato N.;  
 RT "Characterization of Brachyury downstream notochord genes in the Clona  
 RL intestinalis embryo."  
 Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO OTHER SUBUNITS (ALPHA) OF SUCCINYL-COA SYNTHETASE.  
 CC -1- SIMILARITY: TO OTHER SUBUNITS (BETA) OF SUCCINYL-COA SYNTHETASE,  
 CC OF MALATE-COA LIGASE AND TO ATP CITRATE-LYASE.  
 CC EMBL: AB036847; BAB00624.1;  
 DR InterPro: IPR000303; COA\_Ligase.  
 DR InterPro: IPR000901; CPase.  
 DR Pfam: PF00549; Ligase-CoA; 1.  
 DR PROSITE: PS00867; CPase-CoA; 1.  
 DR PROSITE: PS01216; SUCCINYL-COA\_Lig\_1; 1.  
 DR PROSITE: PS03399; SUCCINYL-COA\_Lig\_2; UNKNOWN\_1.  
 DR PROSITE: PS01217; SUCCINYL-COA\_Lig\_3; 1.  
 KW Ligase; Lyase; Phosphorylation.  
 SQ SPOUNCE 1101 AA; 120728 MW; 63CF1B8D285DF49 CRC64;

Query Match  
 Best Local Similarity 45.8%; Score 44; DB 5; Length 1101;  
 Matches 50.0%; Pred. No. 64;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 CNEVDVTTRLE 15  
 Db 787 CSENLTKIRE 798

RESULT 11  
 ID 082667 PRELIMINARY; PRT; 544 AA.  
 AC 082667;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE SPIKE PROTEIN SUBUNIT 1 PRECURSOR (FRAGMENT).  
 GN S.  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-UK/918/67;  
 RX MEDLINE=93298060; PubMed=8390829;  
 RA Cavanagh D., Davis P.J.;  
 RA "Sequence analysis of strains of avian infectious bronchitis  
 RL Arch. Virol. 130:471-476(1992).  
 DR EMBL: X64737; CAA46003.1;  
 DR InterPro: IPR002551; Corona\_SI.  
 DR Pfam: PF01600; Corona\_SI; 1.  
 KW Signal; Membrane.  
 FT CHAIN 1 18  
 FT NON-TER 19 >544  
 FT SEQUENCE 544 AA; 59856 MW; 0A0A2F26344CE0B CRC64;

Query Match  
 Best Local Similarity 44.8%; Score 43; DB 12; Length 544;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 HCNNDVT 10  
 Db 101 HCNNDVT 108

RESULT 12  
 ID 082619 PRELIMINARY; PRT; 545 AA.  
 AC 082619;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE SUPRACE GLYCOPROTEIN (FRAGMENT).  
 GN SPIKE\_SI.  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HOLTE;  
 RA Wang L., Collisson E.W.;  
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L18988; AAB74379.1;  
 DR InterPro: IPR002551; Corona\_SI.  
 DR Pfam: PF01600; Corona\_SI; 1.  
 FT NON-TER 545 545  
 SQ SEQUENCE 545 AA; 60125 MW; C40BF969ABA9647 CRC64;

Query Match  
 44.8%; Score 43; DB 12; Length 545;

Best Local Similarity 75.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 HCNFNDVT 10  
Db 101 HCNFDDIT 108

RESULT 13  
QY9VE6 PRELIMINARY; PRT: 739 AA.  
AC QY9VE6; 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
DE CG3136 PROTEIN.  
GN CG3136.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Egharshian A.E., Garg N.S., Gelbart W.M., Glasser W.,  
RA Foster C., Garfield A.E., Garg N.S., Gelbart W.M., Glasser W.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Sidenkham I., Simpson M., Skupski M.P., Smith T.,  
RA Splier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003787; AAF57345.1; -  
DR Flybase: FBgn0033010; CG3136.  
DR InterPro: IPR001871; bZIP.  
DR Pfam: PF00170; bZIP.1.  
DR SMART: SM00338; BRLZ; 1.  
SO SEQUENCE 739 AA; 84551 MW; 63067E2E08F84F7 CRC64;

Query Match 44.8%; Score 43; DB 5; Length 739;

Best Local Similarity 40.0%; Pred. No. 63;  
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
QY 2 EHCNFDVTRLREN 16  
Db 374 KHCNNTTATIKKN 388

RESULT 14  
QY0950 PRELIMINARY; PRT: 96 AA.  
AC QY0950; 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
DE ENVELOPE GLYCOPROTEIN VIA2 REGION (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OC NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=065;  
RA Cornelissen M., Goudsmit J.;  
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=065;  
RA MCEVILLY M.M.;  
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U13544; AAA73733.1; -  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
FT NON\_TER 96 96  
SQ SEQUENCE 96 AA; 10776 MW; 9B9443D3CD7F43A8 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 96;  
Best Local Similarity 43.8%; Pred. No. 10;  
Matches 7; Conservative 6; Mismatches 2; Indels 1; Gaps 1;  
QY 2 EHCNFDVTRLREN 17  
Db 43 EYCSFN-ITTEIRDQ 57

RESULT 15  
ID P90081 PRELIMINARY; PRT: 240 AA.  
AC P90081; 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OC NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PATIENT A;  
RX MEDLINE=97184515; PubMed=9032317;  
RA McDonald R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
RA Birx D.L., Michael N.L.;  
RT "Evolution of human immunodeficiency virus type 1 env sequence  
variation in patients with diverse rates of disease progression and T-  
cell function.";  
RL J. Virol. 71:1871-1879(1996).  
DR EMBL: U69285; AAC56604.1; -  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.

KW Envelope protein.  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SO SEQUENCE 240 AA; 26720 MW; 32760466F045DAA2 CRC64;

Query Match  
 Best Local Similarity 44.3%; Score 42.5; DB 12; Length 240;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;  
 QY 2 EHCNNDVTTRLRE 15  
 DB 32 KNCSEF-ITTRLRD 44

RESULT 16  
 ID P90082 PRELIMINARY; PRT; 240 AA.  
 AC P90082;  
 DT 01-MAY-1997 (TREMblrel. 03, Created)  
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)  
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT A;  
 RX MEDLINE=97184515; PubMed=9032317;  
 RA McDONALD R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
 RT Birx D.L., Michael N.L.;  
 RT "Evolution of human immunodeficiency virus type 1 env sequence  
 cell function."  
 RL J. Virol. 71:1871-1879(1996).  
 DR EMBL; 069287; AAC56606.1; -;  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW Envelope protein.  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SO SEQUENCE 240 AA; 26756 MW; 30B077048669D793 CRC64;

Query Match  
 Best Local Similarity 44.3%; Score 42.5; DB 12; Length 240;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;  
 QY 2 EHCNNDVTTRLRE 15  
 DB 32 KNCSEF-ITTRLRD 44

RESULT 17  
 ID P90085 PRELIMINARY; PRT; 240 AA.  
 AC P90085;  
 DT 01-MAY-1997 (TREMblrel. 03, Created)  
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)  
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT A;  
 RX MEDLINE=97184515; PubMed=9032317;  
 RA McDONALD R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
 RT Birx D.L., Michael N.L.;

RT "Evolution of human immunodeficiency virus type 1 env sequence  
 cell function."  
 RT Variation in patients with diverse rates of disease progression and T-  
 RL J. Virol. 71:1871-1879(1996).  
 DR EMBL; 069291; AAC56610.1; -;  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW Envelope protein.  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SO SEQUENCE 240 AA; 26729 MW; 4A3C28B659E493B CRC64;

Query Match  
 Best Local Similarity 44.3%; Score 42.5; DB 12; Length 240;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;  
 QY 2 EHCNNDVTTRLRE 15  
 DB 32 KNCSEF-ITTRLRD 44

RESULT 18  
 ID P90080 PRELIMINARY; PRT; 240 AA.  
 AC P90080;  
 DT 01-MAY-1997 (TREMblrel. 03, Created)  
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)  
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT A;  
 RX MEDLINE=97184515; PubMed=9032317;  
 RA McDONALD R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
 RT Birx D.L., Michael N.L.;  
 RT "Evolution of human immunodeficiency virus type 1 env sequence  
 cell function."  
 RL J. Virol. 71:1871-1879(1996).  
 DR EMBL; 069284; AAC56603.1; -;  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW Envelope protein.  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SO SEQUENCE 240 AA; 26747 MW; B85EA6F749552442 CRC64;

Query Match  
 Best Local Similarity 44.3%; Score 42.5; DB 12; Length 240;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;  
 QY 2 EHCNNDVTTRLRE 15  
 DB 32 KNCSEF-ITTRLRD 44

RESULT 19  
 ID P90083 PRELIMINARY; PRT; 240 AA.  
 AC P90083;  
 DT 01-MAY-1997 (TREMblrel. 03, Created)  
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)  
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.



NCBI\_TaxID=11676;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN-PATIENT A;  
 MEDLINE=97184515; PubMed=9032317;  
 McDonald R.A., Mayers D.L., Chung R.C.Y., Wagner K.F., Kim S.,  
 Birx D.L., Michael N.L.;  
 "Evolution of human immunodeficiency virus type 1 env sequence  
 variation in patients with diverse rates of disease progression and T-  
 cell function."  
 J. Virol. 71:1871-1879(1996).  
 EMBL; U69289; AAC56808.1; -;  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 KM Envelope protein.  
 FT NON\_TER 1 1  
 DR 240 240  
 FT SEQUENCE 240 AA; 26747 MW; B85EA6F749552442 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 240;  
 Best Local Similarity 57.1%; Pred. No. 25;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 15  
 Db 32 KNCSPN-ITRLRD 44

RESULT 20  
 ID 09S091 PRELIMINARY; PRT; 313 AA.  
 AC 09S091;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE HYPOHERICAL 36.0 KDA PROTEIN.  
 GN BHO17.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OG Plasmid cp32-7.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,  
 RA White O., Dodson R., Hickey E.K., Gwin M., Peterson J., van-Vugt R.,  
 RA Palmer N., Haft D., Rosa P., Stevenson B.;  
 RT "A bacterial genome in flux: The twelve linear and nine circular  
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease  
 RT spirochete Borrelia burgdorferi."  
 RL Mol. Microbiol. 0:0-0(1999).  
 DR EMBL; AE001579; AAF07613.1; -;  
 KM Hypothetical protein; Plasmid.  
 SQ SEQUENCE 313 AA; 35968 MW; E81A8779733736C5 CRC64;

Query Match 44.3%; Score 42.5; DB 2; Length 313;  
 Best Local Similarity 56.2%; Pred. No. 32;  
 Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 4 CNF--NDVTRLREN 16  
 Db 211 CNFETNDITTEADEN 226

RESULT 21

ID 09IK02 PRELIMINARY; PRT; 679 AA.  
 AC 09IK02;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE TRUNCATED ENVELOPE GLYCOPROTEIN.

ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-546HC-F7;  
 RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
 RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
 RT a patient with HIV dementia: evidence for monocyte trafficking into  
 RT brain."  
 J. Neurovirol. 0:0-0(2000).  
 RL EMBL; AF217165; AAF75507.1; -;  
 DR InterPro: IPR000328; Env\_Gp41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KM Envelope protein.  
 SQ SEQUENCE 679 AA; 76494 MW; 69C6485D07902578 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 679;  
 Best Local Similarity 57.1%; Pred. No. 70;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 15  
 Db 153 KNCSPN-ITRLRD 165

RESULT 22  
 ID 09IK06 PRELIMINARY; PRT; 684 AA.  
 AC 09IK06;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-546HC-B6;  
 RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
 RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
 RT a patient with HIV dementia: evidence for monocyte trafficking into  
 RT brain."  
 RL J. Neurovirol. 0:0-0(2000).  
 DR EMBL; AF217161; AAF75503.1; -;  
 DR InterPro: IPR000328; Env\_Gp41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 SQ SEQUENCE 684 AA; 77266 MW; F9E6F3CCE4D32E10 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 684;  
 Best Local Similarity 57.1%; Pred. No. 70;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRE 15  
 Db 153 KNCSPN-ITRLRD 165

RESULT 23

ID 073350 PRELIMINARY; PRT; 852 AA.  
 AC 073350;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
 DE ENVELOPE GLYCOPROTEIN.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 RX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BU/91/07;  
 RX MEDLINE=96303593; PubMed=8744585.  
 RA Penny M.A., Thomas S.J., Douglas N.W., Ranjbar S., Holmes H.,  
 RA Daniels R.S.;  
 RT "env gene sequences of primary HIV type 1 isolates of subtypes B, C,  
 RT D, E, and F obtained from the World Health Organization Network for  
 RT HIV Isolation and Characterization."  
 RL AIDS Res. Hum. Retroviruses 12:741-747(1996).  
 DR EMBL; U39245; AAB37182.1; -  
 DR InterPro; IPR000328; Env.GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 RW Envelope protein.  
 SQ SEQUENCE 852 AA; 96383 MW; 8F2AC94B5FF7492B CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 852;  
 Best Local Similarity 50.0%; Pred. No. 87;  
 Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;  
 Oy 2 EHCNFDVTRLRNE 17  
 Db 155 KNCSEF-ITTRLDKE 169

RESULT 24  
 O91K05 PRELIMINARY; PRT; 861 AA.  
 AC O91K05;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
 DE 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 RX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=546HC-C5;  
 RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
 RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
 RT brain."  
 RT J. Neurovirol. 0:0-0(2000).  
 DR EMBL; AF217163; AAF75504.1; -  
 DR InterPro; IPR000328; Env.GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 SQ SEQUENCE 861 AA; 97808 MW; 02972C8F2C9464D3 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 861;  
 Best Local Similarity 57.1%; Pred. No. 88;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;  
 Oy 2 EHCNFDVTRLRNE 15  
 Db 153 KNCSEF-ITTRLD 165

RESULT 25  
 O91K04

ID O91K04 PRELIMINARY; PRT; 861 AA.  
 AC O91K04;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
 DE 01-JUN-2001 (TReMBLrel. 17, last annotation update)  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 RX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=546HC-C8;  
 RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
 RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
 RT brain."  
 RT J. Neurovirol. 0:0-0(2000).  
 DR EMBL; AF217163; AAF75505.1; -  
 DR InterPro; IPR000328; Env.GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 SQ SEQUENCE 861 AA; 97763 MW; 0CEB8C39E698D375 CRC64;

Query Match 44.3%; Score 42.5; DB 12; Length 861;  
 Best Local Similarity 57.1%; Pred. No. 88;  
 Matches 8; Conservative 5; Mismatches 0; Indels 1; Gaps 1;  
 Oy 2 EHCNFDVTRLRNE 15  
 Db 153 KNCSEF-ITTRLD 165

Search completed: April 2, 2002, 09:23:52  
 Job time: 196 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 10:43:55 ; Search time 23.72 Seconds

(Without alignments)  
53.088 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVTTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 50 summaries

Database :

A\_Geneseq\_1101.\*  
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                  |
|------------|-------|-------------|--------|-------|------------------------------|
| 1          | 96    | 100.0       | 77     | 20    | AAV27313 Human CD59 protein  |
| 2          | 96    | 100.0       | 102    | 18    | AAW09041 Human membrane att  |
| 3          | 96    | 100.0       | 103    | 12    | AA11426 Human lymphocyte s   |
| 4          | 96    | 100.0       | 103    | 14    | AA132291 Sequence of CD59.   |
| 5          | 96    | 100.0       | 107    | 16    | AA180239 Human membrane att  |
| 6          | 96    | 100.0       | 115    | 21    | AA180376 Human secreted pro  |
| 7          | 96    | 100.0       | 115    | 21    | AA180376 Human secreted pro  |
| 8          | 96    | 100.0       | 127    | 20    | AAV27311 Human CD59 protein  |
| 9          | 96    | 100.0       | 128    | 11    | AA1804704 Sequence of the IF |
| 10         | 96    | 100.0       | 128    | 11    | AA1804704 Human membrane att |
| 11         | 96    | 100.0       | 128    | 11    | AA1804704 Human membrane att |

|    |    |       |     |    |                             |
|----|----|-------|-----|----|-----------------------------|
| 12 | 96 | 100.0 | 128 | 12 | AA11876 Human lymphocyte s  |
| 13 | 96 | 100.0 | 128 | 16 | AA180240 Human membrane att |
| 14 | 96 | 100.0 | 128 | 16 | AA180315 Human CD59. Homo   |
| 15 | 96 | 100.0 | 128 | 18 | AA180315 Human CD59. Homo   |
| 16 | 96 | 100.0 | 128 | 22 | AA180315 Human CD59. Homo   |
| 17 | 96 | 100.0 | 260 | 22 | AA180315 Human CD59. Homo   |
| 18 | 96 | 100.0 | 261 | 22 | AA180315 Human CD59. Homo   |
| 19 | 68 | 70.8  | 58  | 22 | AA180315 Human CD59. Homo   |
| 20 | 68 | 70.8  | 58  | 22 | AA180315 Human CD59. Homo   |
| 21 | 68 | 70.8  | 58  | 22 | AA180315 Human CD59. Homo   |
| 22 | 68 | 70.8  | 58  | 22 | AA180315 Human CD59. Homo   |
| 23 | 63 | 65.6  | 77  | 20 | AA180315 Human CD59. Homo   |
| 24 | 63 | 65.6  | 77  | 20 | AA180315 Human CD59. Homo   |
| 25 | 63 | 65.6  | 77  | 20 | AA180315 Human CD59. Homo   |
| 26 | 63 | 65.6  | 77  | 20 | AA180315 Human CD59. Homo   |
| 27 | 63 | 65.6  | 77  | 20 | AA180315 Human CD59. Homo   |
| 28 | 63 | 65.6  | 77  | 20 | AA180315 Human CD59. Homo   |
| 29 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 30 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 31 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 32 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 33 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 34 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 35 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 36 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 37 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 38 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 39 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 40 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 41 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 42 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 43 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 44 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 45 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 46 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 47 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 48 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 49 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |
| 50 | 49 | 51.0  | 77  | 20 | AA180315 Human CD59. Homo   |

#### ALIGNMENTS

|          |   |                           |
|----------|---|---------------------------|
| RESULT 1 | AAV27313  | standard; peptide: 77 AA. |
| ID       | AAV27313  |                           |
| XX       | AAV27313  |                           |
| AC       | AAV27313  |                           |
| XX       | AAV27313  |                           |
| DT       | 05-NOV-1999   | (first entry)             |
| XX       | 05-NOV-1999   |                           |
| DE       | Human CD59 protein fragment.  |                           |
| XX       | Human CD59 protein fragment.  |                           |
| KW       | CD59 mediated complement; human; Cd59 protein; mimetic;               |                           |
| KW       | tumour therapy; complement-mediated inflammation; immune disorder;    |                           |
| KW       | immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex; |                           |
| KW       | plasma membrane antigen.  |                           |
| XX       | plasma membrane antigen.  |                           |
| OS       | Homo sapiens.   |                           |
| XX       | Homo sapiens.   |                           |
| PN       | MO9940115-A2.   |                           |
| XX       | MO9940115-A2.   |                           |
| PD       | 12-AUG-1999.  |                           |
| XX       | 12-AUG-1999.  |                           |
| PF       | 09-FEB-1999;  | 99WO-US02802.             |
| XX       | 09-FEB-1999;  |                           |
| PR       | 09-FEB-1998;  | 98US-0020393.             |
| XX       | 09-FEB-1998;  |                           |
| PA       | (BLOO-) BLOOD CENT RES FOUND INC.                                     |                           |
| XX       | (BLOO-) BLOOD CENT RES FOUND INC.                                     |                           |
| PA       | (OKLA-) OKLAHOMA MEDICAL RES FOUND.                                   |                           |
| XX       | (OKLA-) OKLAHOMA MEDICAL RES FOUND.                                   |                           |
| PI       | Sims PJ;  |                           |

XX  
DR WPI: 1999-527301/44.  
XX  
PI Compounds modulating CD59 mediated complement activity, for  
PI treatment of, e.g. immunovascularitis  
XX  
PS Disclosure: Fig 8A-B; 75pp; English.  
XX  
CC The invention relates to compounds modulating CD59 mediated complement  
CC activity. It provides (1) molecules structurally mimicking human CD59  
CC amino acid residues 42-58 (region which serves as binding site for CD59  
CC -C9 interactions) when they are in a spatial orientation which can  
CC inhibit the formation of the human C5b-9 complex. These mimetics  
CC specifically bind to amino acid residues 359-384 of human C9. (11)  
CC in a spatial orientation mimicking C9 amino acids 359-384 when they are  
CC complex. Compounds that mimic CD59 can be used to increase CD59  
CC inhibition of C5b-9 complex assembly. This is especially useful in  
CC patients in need of suppression of complement-mediated inflammation, e.g.  
CC arthritis, scleroderma, and diseases such as immunovascularitis, rheumatoid  
CC C5b-9 complex assembly. This is useful in patients in need of promote  
CC activation. The present sequence represents a human CD59 (a plasma membrane  
CC antigen) protein fragment.  
CC  
XX  
SO Sequence 77 AA;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 20; Length 77;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNNDVYTRIRENE 17  
DB 42 fchcnndvtrlirene 58  
|||||

RESULT 2  
AAW09041  
ID AAW09041 standard; Protein; 102 AA.  
XX  
AC AAW09041;  
XX  
DT 23-SEP-1997 (first entry)  
XX  
DE Human membrane attack complex inhibitory factor mutant Asn18Gln.  
XX  
KW Human; membrane attack complex; inhibitory factor; MACIF; mutant;  
KW fibroblast growth factor; mediation; cytolysis; PG12; bFGF; basic;  
XX transplant rejection.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Peptide 1..25 Location/Qualifiers  
FT Peptide 26..102 /label= sig\_peptide  
FT Misc-difference 43 /label= mat\_peptide  
FT /note= "wild type Asn replaced by Gln"  
PN W09700320-A1.  
XX  
PD 03-JAN-1997.  
XX  
PF 13-JUN-1996; 96MO-JP01609.  
XX  
PR 16-JUN-1995; 95JP-0174282.  
XX  
PA (YAMA ) YAMANOUCHI PHARM CO LTD.  
XX

PI Egashira A, Itou K, Masuho Y, Sugita Y, Suzuki H;  
PI Yamaji N, Yasunaga K;  
XX  
DR WPI: 1997-077527/07.  
DR N-PSDB; AAT49584.  
XX  
PT Modified human membrane attack complex inhibitory factor - has  
PT glutamine at position 18 and is an antiinflammatory and complement  
XX blocker, useful for inhibition of transplant hyper-acute rejection  
PS Claim 1; Page 35; 59pp; Japanese.  
XX  
CC The present sequence is the human membrane attack complex  
CC inhibitory factor (MACIF) mutant Asn18Gln (residues 1-77), which  
CC basic fibroblast growth factor (bFGF) by complement. It can be used  
CC to suppress transplanted organ rejection, especially hyperacute  
CC transplant rejection, and the non-lethal effects of complement,  
XX e.g. inflammation.  
XX  
SO Sequence 102 AA;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 18; Length 102;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNNDVYTRIRENE 17  
DB 67 fchcnndvtrlirene 83  
|||||

RESULT 3  
AAR11426  
ID AAR11426 standard; Protein; 103 AA.  
XX  
AC AAR11426;  
XX  
DT 11-JUN-1991 (first entry)  
XX  
DE Human lymphocyte surface antigen.  
XX  
KW Lymphocyte; surface antigen; autoimmune disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN JP03048696-A.  
XX  
PD 01-MAR-1991.  
XX  
PF 14-JUL-1989; 89JP-0183264.  
XX  
PR 14-JUL-1989; 89JP-0183264.  
XX  
PA (TORA ) TORAY IND INC.  
XX  
DR WPI: 1991-106290/15.  
DR N-PSDB; AAQ11251.  
XX  
PT Surface antigen of human lymphocyte - for use in study of human  
PT immune system of 101 residues  
PS Claim 1; Page 1; 12pp; Japanese.  
XX  
CC This protein is similar to the murine lymphocyte antigen Ly 6.  
CC Antibodies raised against the antigen can be used in formulations  
CC for the treatment of autoimmune diseases, cancer and infections.  
CC The corresponding nucleotide sequence does not include an initial  
CC leu codon.  
XX See also AAQ11252-4.  
XX  
SO Sequence 103 AA;

```

Query Match      100.0%; Score 96; DB 12; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTTLRENE 17
   |||
Db 42 fenhcnfdvttlrilrene 58

RESULT 4
AAR32291
ID AAR32291 standard; Protein; 103 AA.
AC AAR32291;
XX
XX 08-JUN-1993 (first entry)
DE Sequence of CD59.
XX
XX Complement mediated attack inhibitor; CD59; transplant rejection; ss.
XX
XX Homo sapiens.
XX
XX WO9302188-A.
XX
XX 04-FEB-1993.
XX
XX 14-JUL-1992; 92MO-US05920.
XX
XX 15-JUL-1991; 91US-0729926.
XX
XX 29-JUN-1992; 92US-0906394.
XX
XX (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX (UYVA ) UNIV YALE.
XX
XX Bell L, Bothwell ALM, Elliot EA, Flavell RA, Madri J;
XX
XX Rollins S, Sims PJ, Squinto S;
XX
XX WPI: 1993-058786/07.
XX
XX N-PSDB; AAQ36708.
XX
XX Genetically engineered mammalian cell for treatment of coronary
XX
XX artery disease - inhibits complement-mediated attack and does not
XX
XX express surface proteins encoded by class I or II major
XX
XX histocompatibility complex genes
XX
XX Claim 31: Page 75; 89pp; English.
XX
XX The inventors claim a cell which contains a gene sequence which
XX
XX encodes protein CD59. CD59 is expressed by the cell and CD59 inhibits
XX
XX complement mediated attack of the cell. The cells fail to elicit T
XX
XX lymphocyte mediated attack or are resistant to complement mediated
XX
XX attack. They can be used to treat patients with immune disorders.
XX
XX Sequence 103 AA:

Query Match      100.0%; Score 96; DB 14; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTTLRENE 17
   |||
Db 42 fenhcnfdvttlrilrene 58

RESULT 5
AAR80239
ID AAR80239 standard; Protein; 107 AA.
AC AAR80239;
XX

```

```

DT 17-APR-1996 (first entry)
XX
XX Human membrane attack complex inhibition factor.
DE
XX
XX MACIF: membrane attack complex inhibition factor; complement system;
XX
XX regulation; activation; final stage; inhibit damage; disease therapy;
XX
XX type II allergy; type III allergy; inflammatory disease treatment;
XX
XX phosphatidylinositol anchor; glycoprotein; diagnosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX
XX Peptide 1..25
XX
XX /label= secretory_signal_sequence
XX
XX /note= "optionally not present, replaced with
XX
XX H or Met"
XX
XX Peptide 101..107
XX
XX /note= "optionally not present"
XX
XX Peptide 102..107
XX
XX /note= "optionally not present"
XX
XX Protein 26..95
XX
XX /note= "core human MACIF protein"
XX
XX Peptide 96..101
XX
XX /label= PI_attachment_signal_sequence
XX
XX Modified-site 101
XX
XX /label= PI_anchor
XX
XX /note= "modified by PI - skeletal structure composed
XX
XX of phospho-ethanolamine, glycan and
XX
XX phosphatidylinositol"
XX
XX EP672683-A1.
XX
XX 20-SEP-1995.
XX
XX 19-APR-1990; 90EP-0200379.
XX
XX 27-OCT-1989; 89JP-0281197.
XX
XX 21-APR-1989; 89JP-0103088.
XX
XX 12-JUL-1989; 89JP-0179933.
XX
XX 06-SEP-1989; 89JP-0230983.
XX
XX 13-SEP-1989; 89JP-0238246.
XX
XX 21-SEP-1989; 89JP-0247818.
XX
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
XX
XX Furuchi K, Ito K, Sugita Y, Takayama M, Tomita M;
XX
XX Yamaji N, Yano S, Yusakawa K, Kasukawa K, Takemoto T;
XX
XX WPI: 1995-321975/42.
XX
XX N-PSDB; AAQ98532.
XX
XX Peptide with human membrane attack complex inhibition factor
XX
XX activity - also DNA and expression vectors used to regulate the
XX
XX complement system in the final stage of complement activation
XX
XX Claim 2; Page 28; 49pp; English.
XX
XX Human membrane attack complex inhibition factor (MACIF) regulates the
XX
XX complement system in the final stage of complement activation, and
XX
XX inhibits damage of human cells and tissues as a result of MAC
XX
XX formation. Naturally occurring human MACIF is a glycoprotein of mol. wt.
XX
XX 18 +/- 1 kDa (by SDS-PAGE) with a phosphatidylinositol (PI) anchor at
XX
XX position 76 (Glu) at the C-terminus. When the gene encoding MACIF is
XX
XX expressed in bacteria, the gene (see AAQ98532) gives a modified human
XX
XX MACIF protein comprising 128 amino acid residues (AAR80240); the PI
XX
XX anchor attachment does not occur in that case.
XX
XX Sequence 107 AA:

Query Match      100.0%; Score 96; DB 16; Length 107;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 FEHCNFNDVTTLRLENE 17  
 Db 67 fehcnfndvttlrleene 83

## RESULT 6

AAAG03764  
 ID AAG03764 standard; Protein; 115 AA.

AC AAG03764;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7845.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

OS Homo sapiens.

PI EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

DR N-PSDB; AAC03770.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 13; SEQ ID 7845; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 were prepared from total human RNAs or polyA+ RNAs derived from 30  
 different tissues. EST sequences usually correspond mainly to the 3'  
 untranslated region (UTR) of the mRNA because they are often obtained  
 from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 those cases where longer cDNA sequences have been obtained, the full 5'  
 UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 ends and can therefore be used to obtain full length cDNAs and genomic  
 DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 chromosome mapping procedures. They are used to obtain upstream  
 regulatory sequences and to design expression and secretion vectors.

SO Sequence 115 AA;

Query Match 100.0%; Score 96; DB 21; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FEHCNFNDVTTLRLENE 17  
 Db 67 fehcnfndvttlrleene 83

## RESULT 7

AAAG03765

ID AAG03765 standard; Protein; 115 AA.

AC AAG03765;

XX 06-OCT-2000 (first entry)

DT Human secreted protein, SEQ ID NO: 7846.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

OS Homo sapiens.

PI EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

DR N-PSDB; AAC03771.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 13; SEQ ID 7846; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 were prepared from total human RNAs or polyA+ RNAs derived from 30  
 different tissues. EST sequences usually correspond mainly to the 3'  
 untranslated region (UTR) of the mRNA because they are often obtained  
 from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 those cases where longer cDNA sequences have been obtained, the full 5'  
 UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 ends and can therefore be used to obtain full length cDNAs and genomic  
 DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 chromosome mapping procedures. They are used to obtain upstream  
 regulatory sequences and to design expression and secretion vectors.

SO Sequence 115 AA;

Query Match 100.0%; Score 96; DB 21; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FEHCNFNDVTTLRLENE 17  
 Db 67 fehcnfndvttlrleene 83

## RESULT 8

AAAG03766

ID AAG03766 standard; Protein; 115 AA.

AC AAG03766;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7847.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.

OS Homo sapiens.

PI EPI033401-A2.



```

XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-0200610.
PE
XX 26-FEB-1999; 99US-0122487.
PR
XX (GENSET ) GENSET.
PA
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI: 2000-500381/45.
DR N-PSDB; AAC03772.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 13: SEQ ID 7847; 71pp + CD-ROM; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30'
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SO Sequence 115 AA:

Query Match 100.0%; Score 96; DB 21; Length 115;
Best Local Similarity 100.0%; Pred. No. 8.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRRLRENE 17
   |||||
Db 67 fehcnfdvtrrlrene 83

RESULT 9
AAV27311
ID AAV27311 standard; peptide: 127 AA.
XX
AC AAV27311;
XX
DT 05-NOV-1999 (first entry)
XX
DE Human CD59 protein sequence.
XX
KW CD59 mediated complement; human; Cd59 protein; C9 protein; mimetic;
KW tumour therapy; complement-mediated inflammation; immune disorder;
KW immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex;
KW plasma membrane antigen.
XX
XX Homo sapiens.
XX OS
XX PN MO9940115-A2.
XX PD 12-AUG-1999.
XX PF 09-FEB-1999; 99WO-US02802.
XX PR 09-FEB-1998; 98US-0020393.
XX PA (BLOO-) BLOOD CENT RES FOUND INC.
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

```

```

XX Sims PJ;
PI
XX WPI: 1999-527301/44.
DR
XX Compounds modulating CD59 mediated complement activity, for
XX treatment of, e.g. immunovascularitis
XX
PS Disclosure; Fig 1A; 75pp; English.
XX
CC The invention relates to compounds modulating CD59 mediated complement
CC activity. It provides (1) molecules structurally mimicking human CD59
CC amino acid residues 42-58 (region which serves as binding site for CD59
CC -C9 interactions) when they are in a spatial orientation which can
CC inhibit the formation of the human C5b-9 complex. These mimetics
CC specifically bind to amino acid residues 359-384 of human C9. (11)
CC molecules structurally mimicking C9 amino acids 359-384 when they are
CC in a spatial orientation which can promote the formation of the C5b-9
CC complex. Compounds that mimic CD59 can be used to increase CD59
CC inhibition of C5b-9 complex assembly. This is especially useful in
CC patients in need of suppression of complement-mediated inflammation, e.g.
CC immune disorders and diseases such as immunovascularitis, rheumatoid
CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote
CC C5b-9 complex assembly. This is useful in patients in need of complement
CC activation. The composition can be administered as an adjunct to tumour
CC therapy. The present sequence represents a human CD59 (a plasma membrane
CC antigen) protein sequence.
XX
SO Sequence 127 AA:

Query Match 100.0%; Score 96; DB 20; Length 127;
Best Local Similarity 100.0%; Pred. No. 9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRRLRENE 17
   |||||
Db 66 fehcnfdvtrrlrene 82

RESULT 10
AAR04704
ID AAR04704 standard; protein: 128 AA.
XX
AC AAR04704;
XX
DT 30-AUG-1990 (first entry)
XX
DE Sequence of the IF5 antigen derived from human cell membrane.
XX
KW Human IF5 antigen; immunodiagnosis; pernicious anemia;
KW rheumatoid arthritis; systemic lupus erythematosus; glomerular nephritis.
XX
XX Homo sapiens.
XX OS
XX PN EP351313-A.
XX PD 17-JAN-1990.
XX PF 11-JUL-1989; 89EP-0401996.
XX PR 23-MAY-1989; 89JP-0172187.
XX PA (MITU ) MITSUBISHI KASEI CORP.
XX PI Okada H, Okada N, Nagami Y, Takahashi K, Takizawa H, Kondo J;
XX WPI: 1990-016630/03.
XX DR N-PSDB; AAQ03116, AAN93318.
XX
PT New glyco:protein IF5 antigen -
PT derived from human cell membrane, inhibits complement-mediated
PT cell membrane damage

```

XX Claim 3; p. 11; 26pp; English.

PS  
CC Typically it is prep'd. from human erythrocytes which are centrifuged and  
CC the cell membrane fraction suspended overnight in buffer contg. 18  
CC n-octyl-beta-D-glucopyranoside (MOG). After centrifugation, the super-  
CC natant is treated with solid (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> to 60% satn. After centrifugation  
CC the pte. is dissolved in buffer contg. 0.18MOC, then dialysed overnight  
CC against the same plus 0.15 M NaCl. 1F5 antigen is a glycoprotein  
CC with a mol. wt. of 20-25 kd. It contains N-glycoside type carbohydrate  
CC chain and phosphatidylinositol. It inhibits complement-mediated cell  
CC monoclone damage. It may be used to generate polyclonal or  
CC the presence of 1F5 antigen on the surface of erythrocytes,  
CC lymphocytes or other cells, thus enabling the diagnosis of diseases,  
CC such as pernicious anaemia, rheumatoid arthritis,  
CC systemic lupus erythematosus and glomerular nephritis. It may be used  
CC to treat disorders in which complement activation is involved. Abs to it  
CC may be used in targeting therapy, such as cancer cell lysis, or  
CC elimination of malignant cells. Residues 27-70 are encoded by  
CC a cDNA fragment of the gene which encodes 1F5 antigen prep'd. by  
CC the polymerase chain reaction (in AAN93318).

SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 11; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFDVYTRLRENE 17  
Db 67 fehcnfdvtrtlrene 83  
|||||

#### RESULT 11

ID AAR07444 standard; protein; 128 AA.

AC AAR07444;

DT 28-JAN-1991 (first entry)

DE Human membrane attack complex inhibition factor (MACIF) gene

DE product.

KW Haemolysis; late complement components.

OS Homo sapiens.

FM Key Location/Qualifiers

FT Protein 26..128 /label=Mature MACIF protein.

FN EP94035-A.

PD 24-OCT-1990.

PE 19-APR-1990; 90EP-0304203.

PR 27-OCT-1989; 89JP-0281197.

PR 21-APR-1989; 89JP-0103088.

PR 12-JUL-1989; 89JP-0179933.

PR 06-SEP-1989; 89JP-0230983.

PR 13-SEP-1989; 89JP-0238246.

PR 21-SEP-1989; 89JP-0247818.

PA (YAMA ) YAMANOUCHI PHARM KK.

PI - Tomita M, Sugita Y, Takemoto T, Furuchi K, Takayama M;

PI Tsuchikawa K, Yano S, Yamaji N, Ito K;

DR WPI; 1990-322496/43.

DR N-PSDB; AA006262.

PT Genes encoding protein with human MACIF activity - also  
PT expression vectors and proteins produced from expression of the  
PT genes.

PS Disclosure; Fig 1; 49pp; English.

CC Gene product may be expressed in large quantities and pure form  
CC from CHO cells, useful for inhibiting the activity of late  
CC complement components ie. haemolysis resulting from MAC formation.

SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 11; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFDVYTRLRENE 17  
Db 67 fehcnfdvtrtlrene 83  
|||||

#### RESULT 12

ID AAR11876 standard; Protein; 128 AA.

AC AAR11876;

DT 24-JUL-1991 (first entry)

DE Human lymphocyte surface antigen precursor.

KW Ly-6; cancer; autoimmune disease; pre A precursor.

OS Homo sapiens.

FM Key Location/Qualifiers

FT Peptide 1..25 /label= Pre A precursor

FT Protein 26..128 /label= Mature surface antigen

PN JP03081297-A.

PD 05-APR-1991.

PF 23-AUG-1989; 89JP-0218183.

PR 23-AUG-1989; 89JP-0218183.

PA (TORA ) TORAY IND INC.

DR WPI; 1991-144848/20.

DR N-PSDB; AAQ11684-6.

PT Surface antigen of human lymphocyte - and gene and antibody  
PT useful for investigation and treatment of autoimmune disease,  
PT cancer, infection, etc.

PS Claim 5; Page 1203; 16pp; Japanese.

CC Clone is derived from the Ly-6 cDNA sequence from a human lymphocyte  
CC cDNA library. The product may be useful in the study and development  
CC of drugs for the treatment of autoimmune diseases and cancer. Probes  
CC may also be developed for the isolation of other gene families.

SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 12; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9e-08;



OY 1 FEHCNFNDVYTRLRENE 17  
 Db 67 fehcnfndvtrllrene 83

RESULT 15  
 AAM26318  
 ID AAM26318 standard; Protein: 128 AA.  
 AC AAM26318;  
 XX  
 DT 17-NOV-1997 (first entry)  
 DE Human CD59.  
 KW CD59; MAC1F; protectin; p18; human; C5b-9 complement inhibitor;  
 KM retrovirus; vector; gene therapy; stem cell.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..25 /label= Sig\_peptide  
 FT Protein 26..128  
 FT /label= Mat\_protein  
 XX  
 PN US5643770-A.  
 PD 01-JUL-1997.  
 XX  
 PF 21-JUL-1994; 94US-0278630.  
 XX  
 PR 21-JUL-1994; 94US-0278630.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Mason JM, Squinto SP;  
 XX  
 DR WPI: 1997-350243/32.  
 DR N-PSDB; AAT84472.  
 XX  
 PT - for transducing cells in body fluids containing complement  
 PS Disclosure; Column 47-50; 32pp; English.  
 XX  
 CC This protein sequence comprises human CD59, an inhibitor of  
 CC complement C5b-9 protein. Claimed retroviral vector particles  
 CC express a complement inhibitor such as CD59, and are thereby  
 CC protected from inactivation upon exposure to body fluids containing  
 CC complement. Also claimed are: (1) a producer cell producing the  
 CC vector particle; and (2) a chimeric retroviral envelope protein  
 CC (see also AAM26325) with at least part of the N-terminal receptor-  
 CC binding domain removed and replaced by a protein domain having a  
 CC complement inhibitor activity. The vector is used in a claimed  
 CC method for transducing cells in the presence of a body fluid  
 CC containing complement, preferably ex vivo, especially for gene  
 CC therapy, e.g. of hereditary or acquired blood disorders by  
 CC transduction of haematopoietic stem cells.  
 XX  
 SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 18; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 9e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FEHCNFNDVYTRLRENE 17  
 Db 67 fehcnfndvtrllrene 83

RESULT 16

AAU00688  
 ID AAU00688 standard; Protein: 128 AA.  
 AC AAU00688;  
 XX  
 DT 07-SEP-2001 (first entry)  
 DE Human CD59 protein.  
 XX  
 KW CTLA4; CD59; human; pig; T-cell activation; C5b-9 inhibitory activity;  
 KM C3 inhibitory activity; cellular immune response; xenotransplantation;  
 KW rat; lagomorph; hare; ungulate; goat; sheep.  
 OS Homo sapiens.  
 XX  
 PN WO200130966-A2.  
 PD 03-MAY-2001.  
 XX  
 PF 21-OCT-2000; 2000WO-US29151.  
 XX  
 PR 22-OCT-1999; 99US-0161186.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Fodor WL, Pizzolato M;  
 XX  
 DR WPI: 2001-300497/31.  
 DR N-PSDB; AAS00682.  
 XX  
 PT Chimeric protein useful for protecting xenotransplanted tissues by  
 PT C5b-9 and/or C3 inhibitory domain, and a T-cell inhibitory domain -  
 XX  
 PS Example 1; Fig 2E(2); 51pp; English.  
 XX  
 CC The sequence represents human CD59 protein which is fused to human or  
 CC porcine CTLA4 protein to form CTLA4-CD59 chimeric protein. Chimeric  
 CC proteins comprising a domain having C5b-9 and/or C3 inhibitory activity  
 CC (e.g. CTLA4) and a domain having T-cell inhibitory activity (e.g. CD59)  
 CC are capable of inhibiting both cellular immune responses and humoral  
 CC immune responses. These polypeptides and their associated nucleic acids  
 CC are useful for protecting pig cells of tissues and organs from both  
 CC humoral and cellular rejection after xenotransplantation into humans. The  
 CC sequences are capable of conferring resistance to humoral and cellular  
 CC mechanisms of immune attack, to protect against human serum complement  
 CC and to inhibit T-cell activation. Transgenic animals (for example,  
 CC rodents, e.g. mouse, rat; lagomorphs, e.g. rabbit, hare; and ungulates,  
 CC e.g. pig, goat, sheep) expressing such a chimeric protein on the surfaces  
 CC of their cells would have a higher chance of survival.  
 XX  
 SQ Sequence 128 AA;

Query Match 100.0%; Score 96; DB 22; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 9e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FEHCNFNDVYTRLRENE 17  
 Db 67 fehcnfndvtrllrene 83

RESULT 17  
 AAU00685  
 ID AAU00685 standard; Protein: 260 AA.  
 AC AAU00685;  
 XX  
 DT 07-SEP-2001 (first entry)  
 DE Human CTLA4-human CD59 chimeric protein.

|    |   |
|----|---|
| XX | CTLA4; CD59; human; pig; T-cell activation; C5b-9 inhibitory activity;    |
| KW | C3 inhibitory activity; cellular immune response; xenotransplantation;    |
| KM | humoral immune response; human serum complement; rodent; mouse; rabbit;   |
| KW | rat; lagomorph; hare; ungulate; goat; sheep; mutant; mutein.              |
| OS | Homo sapiens.   |
| OS | Synthetic.  |
| PN | MO200130966-A2.   |
| PD | 03-MAY-2001.  |
| PF | 21-OCT-2000; 2000MO-US29151.  |
| XX | 22-OCT-1999; 99US-0161186.  |
| XX | (ALEX-) ALEXION PHARM INC.  |
| PA | Fodor WL, Pizzolato M;  |
| PI | WPI: 2001-300497/31.  |
| DR | N-PSDB: AAS00679.   |
| XX | Chimeric protein useful for protecting xenotransplanted tissues by        |
| PT | inhibiting cellular both humoral and immune responses, comprises a        |
| XX | C5b-9 and/or C3 inhibitory domain, and a T-cell inhibitory domain -       |
| XX | Example 1; Fig 2B(2); 51pp: English.                                      |
| XX | The sequence represents a CTLA4-CD59 chimeric protein, formed from human  |
| CC | CTLA4 protein and human CD59 protein. Chimeric proteins comprising a      |
| CC | domain having C5b-9 and/or C3 inhibitory activity (e.g. CTLA4) and a      |
| CC | domain having T-cell inhibitory activity (e.g. CD59) are capable of       |
| CC | inhibiting both cellular immune responses and humoral immune responses.   |
| CC | These polypeptides and their associated nucleic acids are useful for      |
| CC | protecting pig cells of tissues and organs from both humoral and cellular |
| CC | rejection after xenotransplantation into humans. The sequences are        |
| CC | capable of conferring resistance to humoral and cellular mechanisms of    |
| CC | immune attack, to protect against human serum complement and to inhibit   |
| CC | T-cell activation. Transgenic animals (for example, rodents, e.g. mouse,  |
| CC | rat; lagomorphs, e.g. rabbit, hare; and ungulates, e.g. pig, goat, sheep) |
| CC | expressing such a chimeric protein on the surfaces of their cells would   |
| CC | have a higher chance of survival.   |
| XX | Sequence 260 AA;  |
| SO |   |
| XX | Query Match 100.0%; Score 96; DB 22; Length 260;                          |
| XX | Best Local Similarity 100.0%; Pred. NO. 1.9e-07;                          |
| XX | Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0.               |
| QY | 1 FEHCNFDVDTTTLRENE 17  |
| XX |   |
| XX | 199 fchcnfdvdttrlene 215  |
| XX | RESULT 18   |
| XX | AAU00684  |
| XX | ID AAU00684 standard; Protein; 261 AA.                                    |
| XX | AAU00684;   |
| XX | 07-SEP-2001 (first entry)   |
| XX | Porcine CTLA4-human CD59 chimeric protein.                                |
| XX | CTLA4; CD59; human; pig; T-cell activation; C5b-9 inhibitory activity;    |
| KW | C3 inhibitory activity; cellular immune response; xenotransplantation;    |
| KM | humoral immune response; human serum complement; rodent; mouse; rabbit;   |
| KW | rat; lagomorph; hare; ungulate; goat; sheep; mutant; mutein.              |
| OS | Chimeric - Homo sapiens.  |

```
OS Chimeric Sus scrofa.  
XX WO200130966-A2.  
FN  
PD PD3 XX  
PP 03-MAY-2001.  
XX  
PF 21-OCT-2000; 2000WO-US29151.  
XX  
PR 22-OCT-1999; 99US-0161186.  
XX  
PA (ALEX-) ALEXION PHARM INC.  
XX  
PI Fodor WL, Pizzolato M;  
XX WPL: 2001-300497/31.  
DR N-PADB; AAS00678.  
XX  
PT Chimeric protein useful for protecting xenotransplanted tissues by  
inhibiting cellular both humoral and immune responses, comprises a C5b-9 and/or C3 inhibitory domain, and a T-cell inhibitory domain -  
XX Example 3; Fig 2A(2); 51pp; English.
```

The sequence represents a CTLA4-CD5 chimeric protein, formed from porcine CTLA4 protein and human CD5 protein. Chimeric proteins comprising a domain having C5b-9 and/or C3 inhibitory activity (e.g. CTLA4) and a domain having T-cell inhibitory activity (e.g. CD5) are capable of inhibiting both cellular immune responses and humoral immune responses. These polypeptides and their associated nucleic acids are useful for protecting pig cells of tissues and organs from both humoral and cellular rejection after xenotransplantation into humans. The sequences are capable of conferring resistance to humoral and cellular mechanisms of immune attack, to protect against human serum complement and to inhibit T-cell activation. Transgenic animals (for example, rodents, e.g. mouse, rat, lagomorphs, e.g. rabbit, hare; and ungulates, e.g. pig, goat, sheep) expressing such a chimeric protein on the surfaces of their cells would have a higher chance of survival.

```
Sequence      261 AA;
```

Query Match 100.0%; Score 96; DB 22; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0,

OY 1 FEHCNFNDVTRRLRENE 17  
 |||||  
Db 200 fchcnfndvtrlrene 216

Search completed: April 2, 2002, 10:44:27  
Job time: 32 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 93.32 Seconds  
(without alignments)  
50.580 Million cell updates/sec

Title: US-09-020-393B-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVYTRLENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*
- 8: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*
- 9: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*
- 10: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*
- 11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*
- 12: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*
- 13: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*
- 14: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*
- 15: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*
- 16: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*
- 17: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*
- 18: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*
- 19: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*
- 20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*
- 21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*
- 22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*
- 24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                  | Description       |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1          | 96    | 100.0       | 70     | US-08-952-333C-28   | Sequence 28, Appl |
| 2          | 96    | 100.0       | 77     | US-09-020-393B-3    | Sequence 3, Appl  |
| 3          | 96    | 100.0       | 80     | US-08-043-446A-6    | Sequence 6, Appl  |
| 4          | 96    | 100.0       | 103    | US-07-906-394-3     | Sequence 3, Appl  |
| 5          | 96    | 100.0       | 103    | US-08-021-602-3     | Sequence 3, Appl  |
| 6          | 96    | 100.0       | 103    | US-08-086-549-3     | Sequence 3, Appl  |
| 7          | 96    | 100.0       | 122    | US-09-834-366-13482 | Sequence 13482, A |
| 8          | 96    | 100.0       | 122    | US-60-197-873-13482 | Sequence 13482, A |
| 9          | 96    | 100.0       | 127    | US-09-020-393B-1    | Sequence 1, Appl  |

|    |      |       |     |                       |                    |
|----|------|-------|-----|-----------------------|--------------------|
| 10 | 96   | 100.0 | 128 | US-09-834-366-18052   | Sequence 18052, A  |
| 11 | 96   | 100.0 | 128 | US-09-834-366-18056   | Sequence 18056, A  |
| 12 | 96   | 100.0 | 128 | US-09-834-366-18064   | Sequence 18064, A  |
| 13 | 96   | 100.0 | 128 | US-60-197-873-18052   | Sequence 18052, A  |
| 14 | 96   | 100.0 | 128 | US-60-197-873-18056   | Sequence 18056, A  |
| 15 | 96   | 100.0 | 128 | US-60-197-873-18064   | Sequence 18064, A  |
| 16 | 89   | 92.7  | 78  | US-60-160-203-4361    | Sequence 4361, Ap  |
| 17 | 89   | 92.7  | 78  | US-60-169-840-6378    | Sequence 6378, Ap  |
| 18 | 72.5 | 75.5  | 73  | US-60-160-189-7836    | Sequence 7836, Ap  |
| 19 | 72.5 | 75.5  | 73  | US-60-169-867-5425    | Sequence 5425, Ap  |
| 20 | 68   | 70.8  | 58  | PCT-US01-00663-30608  | Sequence 30608, A  |
| 21 | 68   | 70.8  | 58  | US-09-864-761-37475   | Sequence 37475, A  |
| 22 | 68   | 70.8  | 58  | US-60-236-358-19739   | Sequence 19739, A  |
| 23 | 68   | 70.8  | 75  | US-09-020-393B-5      | Sequence 5, Appl   |
| 24 | 63   | 65.6  | 77  | US-09-020-393B-2      | Sequence 9833, Ap  |
| 25 | 52   | 54.2  | 211 | PCT-US01-14827-9833   | Sequence 9741, Ap  |
| 26 | 51   | 53.1  | 47  | US-60-160-203-5893    | Sequence 5893, Ap  |
| 27 | 51   | 53.1  | 47  | US-60-169-840-8842    | Sequence 8842, Ap  |
| 28 | 51   | 53.1  | 47  | US-60-169-867-7696    | Sequence 7696, Ap  |
| 29 | 51   | 53.1  | 47  | US-09-020-393B-6      | Sequence 11, Appl  |
| 30 | 49   | 51.0  | 77  | US-09-020-393B-11     | Sequence 2, Appl   |
| 31 | 49   | 51.0  | 77  | US-09-020-393B-2      | Sequence 7, Appl   |
| 32 | 47   | 49.0  | 77  | US-60-196-174-851     | Sequence 851, App  |
| 33 | 46   | 47.9  | 50  | US-60-196-1174-1174   | Sequence 1174, App |
| 34 | 46   | 47.9  | 50  | US-60-234-446-673     | Sequence 673, App  |
| 35 | 46   | 47.9  | 52  | US-60-196-190-1173    | Sequence 1173, Ap  |
| 36 | 46   | 47.9  | 57  | US-09-252-991A-29464  | Sequence 29464, A  |
| 37 | 46   | 47.9  | 294 | US-09-134-001C-4972   | Sequence 4972, Ap  |
| 38 | 46   | 47.9  | 290 | US-09-450-969-4403    | Sequence 4403, Ap  |
| 39 | 44   | 45.8  | 290 | US-09-188-979-35      | Sequence 35, Appl  |
| 40 | 44   | 45.8  | 15  | US-60-191-681-4548    | Sequence 4548, Ap  |
| 41 | 44   | 45.8  | 15  | US-09-270-767-58029   | Sequence 58029, A  |
| 42 | 43   | 44.8  | 739 | US-09-270-847B-183630 | Sequence 183630, A |
| 43 | 43   | 44.8  | 203 | US-09-248-796-21597   | Sequence 21597, A  |
| 44 | 43   | 44.8  | 203 | US-09-248-796-14911   | Sequence 14911, A  |
| 45 | 42   | 43.8  | 404 | US-09-248-796-14912   | Sequence 14912, A  |
| 46 | 42   | 43.8  | 415 | US-09-270-767-42709   | Sequence 42709, A  |
| 47 | 42   | 43.8  | 16  |                       |                    |
| 48 | 42   | 43.8  |     |                       |                    |
| 49 | 42   | 43.8  |     |                       |                    |
| 50 | 42   | 43.8  |     |                       |                    |

#### ALIGNMENTS

RESULT 1

US-08-952-333C-28

Sequence 28, Application US/08952333C

GENERAL INFORMATION:

APPLICANT: Brakenhoff, Rudolf Henrikus; Van Dongen, Augustina

APPLICANT: Antonius Maria Sylvester

TITLE OF INVENTION: Methods for detection and therapy....eptopes

TITLE OF INVENTION: and antibodies

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Frischaut Holte Goodman Langer and Chick

STREET: 767 Third Avenue

City: New York

STATE: New York

COUNTRY: USA

ZIP: 10017

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/952,333C

FILING DATE: 17 FEBRUARY 1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/NL95/00169

FILING DATE: 10 MAY 1995

ATTORNEY/AGENT INFORMATION:  
NAME: Barth, Richard S  
REGISTRATION NUMBER: 28180  
REFERENCE/DOCKET NUMBER: 970686  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 319-4900  
TELEFAX: (212) 319-5101  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 70 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: protein  
ORGANISM: HCD59 fragment of figure 5  
US-08-952-333C-28

Query Match 100.0%; Score 96; DB 13; Length 70;  
Best Local Similarity 100.0%; Pred. No. 7.2e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 42 FEHCNFDVTTLRRENE 58  
|||||

RESULT 2  
US-09-020-393b-3  
Sequence 3, Application US/09020393b  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
TITLE OF INVENTION: Compositions and Methods to Inhibit the  
NUMBER OF INVENTION: C5b-9 Complex of Complement  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESS: Patrea L. Padst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
STREET: St.  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,393b  
FILING DATE: 03-FEB-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Padst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF 170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
US-09-020-393b-3

Query Match 100.0%; Score 96; DB 14; Length 77;  
Best Local Similarity 100.0%; Pred. No. 8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 42 FEHCNFDVTTLRRENE 58  
|||||

RESULT 3  
US-08-043-446A-6  
Sequence 6, Application US/08043446A  
GENERAL INFORMATION:  
APPLICANT: Boyd, Richard L.  
APPLICANT: Godfrey, Dale I.  
APPLICANT: MacNeil, Ian A.  
APPLICANT: Zlotnik, Albert  
TITLE OF INVENTION: MAMMALIAN DEVELOPMENTAL ANTIGENS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/043,446A  
FILING DATE: 05-APR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0375  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-043-446A-6

Query Match 100.0%; Score 96; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17  
Db 51 FEHCNFDVTTLRRENE 67  
|||||

RESULT 4  
US-07-906-394-3  
Sequence 3, Application US/07906394  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
APPLICANT: Elliott, Eileen A.  
APPLICANT: Flavell, Richard A.  
APPLICANT: Madril, Joseph  
APPLICANT: Rollins, Scott  
APPLICANT: Bell, Leonard  
APPLICANT: Squinto, Stephen





REFERENCE/DOCKET NUMBER: OMRF135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: CD59  
US-08-086-549-3

Query Match  
Best Local Similarity 100.0%; Score 96; DB 4; Length 103;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17  
|||||  
DB 42 FEHCNENDVTTRLRENE 58

RESULT 7  
US-09-834-366-13482  
Sequence 13482, Application US/09834366  
GENERAL INFORMATION:  
APPLICANT: Bejani, Stephane  
APPLICANT: Tanaka, Hiroaki  
APPLICANT: Dumas Milne Edwards, Jean Baptiste  
APPLICANT: Jobert, Severin  
APPLICANT: Giordano, Jean-Yves  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: 81.US2.REG  
CURRENT APPLICATION NUMBER: US/09/834,366  
PRIOR APPLICATION NUMBER: 2001-04-13  
PRIOR FILING DATE: 2000-04-18  
NUMBER OF SEQ ID NOS: 52153  
SOFTWARE: Patent.pm  
SEQ ID NO 13482  
LENGTH: 122  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -20...-1  
US-09-834-366-13482

Query Match  
Best Local Similarity 100.0%; Score 96; DB 22; Length 122;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17  
|||||  
DB 67 FEHCNENDVTTRLRENE 83

RESULT 8  
US-60-197-873-13482  
Sequence 13482, Application US/60197873  
GENERAL INFORMATION:  
APPLICANT: Bejani, Stephane  
APPLICANT: Tanaka, Hiroaki  
APPLICANT: Dumas Milne Edwards, Jean Baptiste

APPLICANT: Jobert, Severin  
APPLICANT: Giordano, Jean-Yves  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: 81.US1.PRO  
CURRENT APPLICATION NUMBER: US/60/197,873  
CURRENT FILING DATE: 2000-04-18  
NUMBER OF SEQ ID NOS: 52153  
SOFTWARE: Patent.pm  
SEQ ID NO 13482  
LENGTH: 122  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -20...-1  
US-60-197-873-13482

Query Match  
Best Local Similarity 100.0%; Score 96; DB 24; Length 122;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17  
|||||  
DB 67 FEHCNENDVTTRLRENE 83

RESULT 9  
US-09-020-393b-1  
Sequence 1, Application US/09020393B  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
TITLE OF INVENTION: Compositions and Methods to Inhibit the  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,393B  
FILING DATE: 03-FEB-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF 170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
US-09-020-393b-1

Query Match 100.0%; Score 96; DB 14; Length 127;

|    | Best Local Similarity | 100.0%; | Pred. No. 1.4e-07; |               |
|----|-----------------------|---------|--------------------|---------------|
|    | Matches               | 17;     | Conservative       | 0; Mismatches |
|    |                       |         |                    | 0; Indels     |
|    |                       |         |                    | 0; Gaps       |
| QY | 1 FEHCNFNDVYTRLRENE   | 17      |                    |               |
|    |                       |         |                    |               |
|    |                       |         |                    |               |
|    |                       |         |                    |               |
| Db | 66 FEHCNFNDVYTRLRENE  | 82      |                    |               |

```

RESULT 10
US-09-834-366-18052
? Sequence 18052, Application US/09834366
? GENERAL INFORMATION:
? APPLICANT: Bejantin, Stephane
? APPLICANT: Tanaka, Hiroaki
? APPLICANT: Dumas, Milne Edwards, Jean Baptiste
? APPLICANT: Jobert, Severin
? APPLICANT: Giordano, Jean-Yves
? TITLE OF INVENTION: ESTs and Encoded Human Proteins
? FILE REFERENCE: 81.US2.REG
? CURRENT APPLICATION NUMBER: US/09/834,366
? CURRENT FILING DATE: 2001-04-13
? PRIOR APPLICATION NUMBER: US 60/197,873
? PRIOR FILING DATE: 2000-04-18
? NUMBER OF SEQ ID NOS: 52153
? SOFTWARE: PatentLm
? SEQ ID NO 18052
? LENGTH: 128
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SIGNAL
? LOCATION: -20...-1
US-09-834-366-18052

```

|                          |        |                   |           |             |
|--------------------------|--------|-------------------|-----------|-------------|
| Query Match              | 100.0% | Score 96;         | DB 22;    | Length 128; |
| Best Local Similarity    | 100.0% | Pred. No. 14e-07; |           |             |
| Matches 17; conservative | 0;     | mismatches 0;     | indels 0; | gaps 0;     |
| QY                       | 1      | FEHCNFDVYTR       | LENE      | 17          |
|                          |        |                   |           |             |
|                          |        |                   |           |             |
|                          |        |                   |           |             |
|                          |        |                   |           |             |
|                          |        |                   |           |             |
| Db                       | 67     | FEHCNFDVYTR       | LENE      | 83          |

```

RESULT 11
US-09-834-366-18056
; Sequence 18056, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas, Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81,US,Reg
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: \ 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197, 873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 18056
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-09-834-366-18056

```

|                       |         |                    |        |             |
|-----------------------|---------|--------------------|--------|-------------|
| Query Match           | 100.0%; | Score 96;          | DB 22; | Length 128; |
| Best Local Similarity | 100.0%; | Pred. No. 1.4e-07; |        |             |

|         |                           |               |           |         |
|---------|---------------------------|---------------|-----------|---------|
| Matches | 17; Conservative          | 0; Mismatches | 0; Indels | 0; Gaps |
| Qy      | 1 FEHCNFNDVYTRLENE 17<br> |               |           |         |
| Db      | 67 FEHCNFNDVYTRLENE 83    |               |           |         |

```

RESULT 12
US-09-834-366-18064
; Sequence 18064, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejantin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Mline Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Ests and Encoded Human Proteins
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patentl.pm
; SEQ ID NO 18064
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
; US-09-834-366-18064

```

|         | Query Match | Similarity   | Score  | DB                | Length |
|---------|-------------|--------------|--------|-------------------|--------|
|         | Best Local  | Similarity   | 100.0% | Pred. No. 1,4e-07 | 128;   |
| Matches | 17;         | Conservative | 0;     | Mismatches        | 0;     |
|         |             |              |        | Indels            | 0;     |
|         |             |              |        | Gaps              | 0;     |
| QY      | 1           | FEHCNFDVYTR  | RENE   | 17                |        |
|         |             |              |        |                   |        |
| DB      | 67          | FEHCNFDVYTR  | RENE   | 83                |        |

```

RESULT 13
US-60-197-873-18052
: Sequence 18052, Application US/60197873
: GENERAL INFORMATION:
: APPLICANT: Bejanin, Stephane
: APPLICANT: Tanaka, Hiroaki
: APPLICANT: Dumas Milne Edwards, Jean Baptiste
: APPLICANT: Jobert, Severin
: APPLICANT: Giordano, Jean-Yves
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: 81.us1 PRO
: CURRENT APPLICATION NUMBER: US/60/197,873
: CURRENT FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 52153
: SOFTWARE: Patent.pm
: SEQ ID NO 18052
: LENGTH: 128
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: -20...-1
: US-60-197-873-18052

```

|                       |                 |                    |           |             |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match           | 100.0%;         | Score 96;          | DB 24;    | Length 128; |
| Best Local Similarity | 100.0%;         | Pred. No. 1,4e-07; |           |             |
| Matches 17;           | Conservative 0; | Mismatches 0;      | Indels 0; | Gaps 0;     |

Db 67 FEHCNFDVTTLRRENE 83

RESULT 14

```
US-60-197-873-18056
; Sequence 18056, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejani, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 18056
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-60-197-873-18056
```

Query Match

Best Local Similarity 100.0%; Score 96; DB 24; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17

Db 67 FEHCNFDVTTLRRENE 83

RESULT 15

```
US-60-197-873-18064
; Sequence 18064, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejani, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 18064
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-60-197-873-18064
```

Query Match

Best Local Similarity 100.0%; Score 96; DB 24; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17

Db 67 FEHCNFDVTTLRRENE 83

RESULT 16

US-60-160-203-4361

```
; Sequence 4361, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4361
; LENGTH: 78
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(78)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-203-4361
```

Query Match

Best Local Similarity 92.7%; Score 89; DB 24; Length 78;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17

Db 17 FEHCNFDVTTLRRENE 33

RESULT 17

```
US-60-169-840-6378
; Sequence 6378, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6378
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(78)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-840-6378
```

Query Match

Best Local Similarity 92.7%; Score 89; DB 24; Length 78;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTLRRENE 17

Db 17 FEHCNFDVTTLRRENE 33

RESULT 18

```
US-60-160-189-7836
; Sequence 7836, Application US/60160189
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; AND USES THEREOF
```

```
; FILE REFERENCE: CL0000112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 7836
; LENGTH: 73
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(73)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-189-7836

Query Match          75.5%; Score 72.5; DB 24; Length 73;
Best Local Similarity 83.3%; Pred. No. 0.00056;
Matches 15; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 1 FEHCNF-NDVTTRLRENE 17
    ||||| |||||
Db 11 FEHCNFRXYVTRLRENE 28

RESULT 19
US-60-169-867-5425
; Sequence 5425, Application US/60169867
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: C1000160
; CURRENT APPLICATION NUMBER: US/60/169,867
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 8230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 5425
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(73)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-867-5425

Query Match          75.5%; Score 72.5; DB 24; Length 73;
Best Local Similarity 83.3%; Pred. No. 0.00056;
Matches 15; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 1 FEHCNF-NDVTTRLRENE 17
    ||||| |||||
Db 11 FEHCNFRXYVTRLRENE 28

RESULT 20
PCT-US01-00663-30608
; Sequence 30608, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
```

```
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO: 30608
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049629.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16
; OTHER INFORMATION: EST HUMAN HIT: A0139072.1, EVALU6 7.00e-29
; OTHER INFORMATION: SWISSPROT HIT: P13987, EVALU6 6.00e-30
PCT-US01-00663-30608

Query Match          70.8%; Score 68; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 NENDVTTRLRENE 17
    ||||| |||||
Db 1 NENDVTTRLRENE 13

RESULT 21
US-09-864-761-37475
; Sequence 37475, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
```

PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 37475  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049629.14  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 16  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EST\_HUMAN HIT: AU139072.1, EVALUE 7.00e-29  
US-09-864-761-37475

Query Match  
Best Local Similarity 70.8%; Score 68; DB 22; Length 58;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 NFNDVTTTRLRENE 17  
Db 1 NFNDVTTTRLRENE 13

RESULT 22  
US-60-236-359-19739  
Sequence 19739, Application US/60236359  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR  
FILE REFERENCE: MDHMRP-4P  
CURRENT FILING DATE: 2000-09-27  
CURRENT APPLICATION NUMBER: US/60/236,359  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 21709  
SOFTWARE: Molecular Dynamics Sequence Listing Engine  
SEQ ID NO 19739  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049629.14  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6  
OTHER INFORMATION: EST\_HUMAN HIT: AW374176.1, EVALUE 4.00e-29  
US-60-236-359-19739

Query Match  
Best Local Similarity 70.8%; Score 68; DB 24; Length 58;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 NFNDVTTTRLRENE 17  
Db 1 NFNDVTTTRLRENE 13

RESULT 23  
US-09-020-393B-4  
Sequence 4, Application US/09020393B  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
TITLE OF INVENTION: Compositions and Methods to Inhibit the  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Padst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 03-FEB-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Padst, Patrea L.  
REGISTRATION NUMBER: 31,284  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Badoon  
US-09-020-393B-4

Query Match  
Best Local Similarity 70.8%; Score 68; DB 14; Length 75;  
Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 FNCNDVTTTRLRENE 17  
Db 42 FNCNDVTTTRLRENE 58

RESULT 24

```
US-09-020-393B-5
; Sequence 5, Application US/09020393B
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: Compositions and Methods to Inhibit the
; TITLE OF INVENTION: C5b-9 Complex of Complement
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 W. Peachtree
; STREET: St.
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-1450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,393B
; FILING DATE: 03-FEB-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR# 170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: African green monkey
;
US-09-020-393B-5

Query Match          65.6%; Score 63; DB 14; Length 77;
Best Local Similarity 58.8%; Pred. No. 0.022;
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRLRENE 17
   1 :|||||:|:|:|
   Db 42 FANCFNDISTLKESE 58

RESULT 25
PCT-US01-14827-9833
; Sequence 9833, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 9833
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (43)..(52)
```

```
; OTHER INFORMATION: Ly-6 / u-PAR domain proteins domain identified by eMATRIX,
; OTHER INFORMATION: accession number BL00983A, p-value=7.480e-09, raw score of 5.8
; NAME/KEY: DOMAIN
; LOCATION: (32)..(160)
; OTHER INFORMATION: u-PAR/Ly-6 domain identified by Pnam, accession name
; NAME/KEY: misc_feature
; LOCATION: (1)..(211)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-14827-9833
```

```
Query Match          54.2%; Score 52; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNEND 8
   1 :|||||
   Db 103 FEHCNEND 110
```

```
Search completed: April 2, 2002, 09:22:42
Job time: 171 sec
```

Tue Apr 2 09:51:56 2002

us-09-020-393b-3\_copy\_42\_58.open.rpm



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 16.47 Seconds  
(without alignments)  
73.774 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRLRENE 17

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 348507 seqs, 71473746 residues  
Total number of hits satisfying chosen parameters: 348507

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                            |
|------------|-------|-------------|--------|-------|--|
| 1          | 96    | 100.0       | 70     | 6     | US-09-612-314-39 Sequence 39, Appl     |
| 2          | 96    | 100.0       | 71     | 6     | US-09-612-314-42 Sequence 42, Appl     |
| 3          | 96    | 100.0       | 77     | 6     | US-09-612-314-37 Sequence 37, Appl     |
| 4          | 96    | 100.0       | 82     | 6     | US-09-612-314-40 Sequence 40, Appl     |
| 5          | 96    | 100.0       | 83     | 6     | US-09-612-314-41 Sequence 41, Appl     |
| 6          | 96    | 100.0       | 88     | 6     | US-09-612-314-45 Sequence 45, Appl     |
| 7          | 96    | 100.0       | 99     | 6     | US-09-612-314-43 Sequence 43, Appl     |
| 8          | 96    | 100.0       | 100    | 6     | US-09-612-314-44 Sequence 44, Appl     |
| 9          | 96    | 100.0       | 128    | 6     | US-09-673-032-20 Sequence 20, Appl     |
| 10         | 96    | 100.0       | 128    | 6     | US-09-614-150-5766 Sequence 5766, Ap   |
| 11         | 41.5  | 43.2        | 236    | 5     | US-09-991-630-12 Sequence 12, Appl     |
| 12         | 41.5  | 43.2        | 236    | 5     | US-09-991-630-15 Sequence 15, Appl     |
| 13         | 41.5  | 43.2        | 236    | 6     | US-09-905-129-12 Sequence 12, Appl     |
| 14         | 41.5  | 43.2        | 236    | 6     | US-09-905-129-15 Sequence 15, Appl     |
| 15         | 40    | 41.7        | 220    | 6     | US-09-708-427-16838 Sequence 16838, A  |
| 16         | 40    | 41.7        | 240    | 6     | US-09-897-516-6395 Sequence 6395, Ap   |
| 17         | 40    | 41.7        | 244    | 6     | US-09-708-427-16837 Sequence 16837, A  |
| 18         | 40    | 41.7        | 244    | 8     | US-60-336-049-36 Sequence 36, Appl     |
| 19         | 40    | 41.7        | 262    | 7     | US-10-092-302-711 Sequence 711, App    |
| 20         | 40    | 41.7        | 262    | 7     | US-10-092-302-951 Sequence 951, App    |
| 21         | 40    | 41.7        | 272    | 6     | US-09-708-427-55600 Sequence 56000, A  |
| 22         | 40    | 41.7        | 299    | 6     | US-09-708-427-55999 Sequence 55999, A  |
| 23         | 40    | 41.7        | 394    | 1     | PCT-US02-03987-15325 Sequence 15325, A |
| 24         | 40    | 41.7        | 394    | 7     | US-10-032-585-7381 Sequence 7381, Ap   |
| 25         | 40    | 41.7        | 394    | 7     | US-10-072-851-15325 Sequence 15325, A  |

|    |    |      |     |   |                                       |
|----|----|------|-----|---|---------------------------------------|
| 26 | 40 | 41.7 | 403 | 6 | US-09-708-427-5395 Sequence 5395, Ap  |
| 27 | 40 | 41.7 | 429 | 6 | US-09-708-427-5394 Sequence 5394, Ap  |
| 28 | 40 | 41.7 | 497 | 6 | US-09-708-427-5393 Sequence 5393, Ap  |
| 29 | 40 | 41.7 | 741 | 6 | US-09-614-150-40656 Sequence 40656, A |
| 30 | 40 | 41.7 | 901 | 6 | US-09-614-150-2010 Sequence 2010, Ap  |
| 31 | 39 | 40.6 | 121 | 6 | US-09-897-516-5831 Sequence 5831, Ap  |
| 32 | 39 | 40.6 | 123 | 6 | US-09-673-032-1 Sequence 1, Appl1     |
| 33 | 39 | 40.6 | 135 | 6 | US-09-897-516-4960 Sequence 4960, Ap  |
| 34 | 39 | 40.6 | 191 | 8 | US-60-361-742-25 Sequence 25, Appl    |
| 35 | 39 | 40.6 | 210 | 5 | US-09-302-626B-169 Sequence 169, Appl |
| 36 | 39 | 40.6 | 213 | 6 | US-09-897-516-7509 Sequence 7509, Ap  |
| 37 | 39 | 40.6 | 222 | 5 | US-09-302-626B-174 Sequence 174, App  |
| 38 | 39 | 40.6 | 240 | 5 | US-09-302-626B-170 Sequence 170, Appl |
| 39 | 39 | 40.6 | 260 | 5 | US-09-589-987-12 Sequence 12, Appl    |
| 40 | 39 | 40.6 | 260 | 5 | US-09-589-987-12 Sequence 12, Appl    |
| 41 | 39 | 40.6 | 260 | 5 | US-09-618-301-12 Sequence 12, Appl    |
| 42 | 39 | 40.6 | 260 | 5 | US-09-618-301-12 Sequence 12, Appl    |
| 43 | 39 | 40.6 | 268 | 7 | US-10-029-386-33039 Sequence 33039, A |
| 44 | 39 | 40.6 | 374 | 6 | US-09-897-516-7828 Sequence 7828, Ap  |
| 45 | 39 | 40.6 | 380 | 6 | US-09-760-446A-1772 Sequence 1772, Ap |
| 46 | 39 | 40.6 | 571 | 5 | US-09-302-626B-26 Sequence 26, Appl   |
| 47 | 39 | 40.6 | 672 | 7 | US-10-010-920-31 Sequence 31, Appl    |
| 48 | 39 | 40.6 | 672 | 8 | US-60-338-287-31 Sequence 31, Appl    |
| 49 | 39 | 40.6 | 680 | 5 | US-09-302-626B-171 Sequence 171, Ap   |
| 50 | 39 | 40.6 | 701 | 6 | US-09-614-150-2730 Sequence 2730, Ap  |

## ALIGNMENTS

```
RESULT 1
US-09-612-314-39
; Sequence 39, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APM634
US-09-612-314-39

Query Match      100.0%; Score 96; DB 6; Length 70;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTRLRENE 17
Db 42 FEHCNFNDVTRLRENE 58

RESULT 2
US-09-612-314-42
; Sequence 42, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
```

```

; APPLICANT: DODD, Ian
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2061
US-09-612-314-42
```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 71;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLRNE 17
Db 42 FEHCNENDVYTRLRNE 58
```

```

RESULT 3
US-09-612-314-37
; Sequence 37, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT631
US-09-612-314-37
```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 77;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLRNE 17
Db 42 FEHCNENDVYTRLRNE 58
```

```

RESULT 4
US-09-612-314-40
```

```

; Sequence 40, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2060
US-09-612-314-40
```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 82;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLRNE 17
Db 42 FEHCNENDVYTRLRNE 58
```

```

RESULT 5
US-09-612-314-41
; Sequence 41, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT635
US-09-612-314-41
```

```

Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 83;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FEHCNENDVYTRLRNE 17
Db 43 FEHCNENDVYTRLRNE 59
```

```
RESULT 6
US-09-612-314-45
; Sequence 45, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT FILING DATE: US/09/612,314
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (71)..(72)
; OTHER INFORMATION: Residues 1-71 are a first polypeptide chain and
; OTHER INFORMATION: Residues 72-88 are a second polypeptide chain
; OTHER INFORMATION: linked by a disulphide bond formed between the
; OTHER INFORMATION: cysteines at positions 71 and 72
; NAME/KEY: UNSURE
; LOCATION: (88)
; OTHER INFORMATION: The C-terminal glycine is linked to an
; OTHER INFORMATION: NH-[Myristoyl] group
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2065
US-09-612-314-45

Query Match          100.0%; Score 96; DB 6; Length 88;
Best Local Similarity 100.0%; Pred. NO. 3.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFDVYTRLRENE 17
   |||||
DB 42 FEHCNFDVYTRLRENE 58

RESULT 7
US-09-612-314-43
; Sequence 43, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT FILING DATE: US/09/612,314
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (82)..(83)
; OTHER INFORMATION: Residues 1-82 are a first polypeptide chain and
; OTHER INFORMATION: Residues 83-99 are a second polypeptide chain
; OTHER INFORMATION: linked by a disulphide bond formed between the
; OTHER INFORMATION: cysteines at positions 82 and 83
; NAME/KEY: UNSURE
; LOCATION: (99)
; OTHER INFORMATION: The C-terminal glycine is linked to an
; OTHER INFORMATION: NH-[Myristoyl] group
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2062
US-09-612-314-43

Query Match          100.0%; Score 96; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. NO. 4.4e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFDVYTRLRENE 17
   |||||
DB 42 FEHCNFDVYTRLRENE 58

RESULT 8
US-09-612-314-44
; Sequence 44, Application US/09612314
; GENERAL INFORMATION:
; APPLICANT: SMITH, Richard, Anthony, Godwin
; APPLICANT: DODD, Ian
; APPLICANT: MOSSAKOWSKA, Danuta, Ewa, Irena
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT FILING DATE: US/09/612,314
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (83)..(84)
; OTHER INFORMATION: Residues 1-83 are a first polypeptide chain and
; OTHER INFORMATION: Residues 84-100 are a second polypeptide chain
; OTHER INFORMATION: linked by a disulphide bond formed between the
; OTHER INFORMATION: cysteines at positions 83 and 84
; NAME/KEY: UNSURE
; LOCATION: (100)
; OTHER INFORMATION: The C-terminal glycine is linked to an
; OTHER INFORMATION: NH-[Myristoyl] group
; OTHER INFORMATION: Description of Artificial Sequence: Protein APT2063
US-09-612-314-44

Query Match          100.0%; Score 96; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. NO. 4.5e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNFDVYTRLRENE 17
   |||||
DB 43 FEHCNFDVYTRLRENE 59

RESULT 9
```

```
US-09-673-032-20
; Sequence 20, Application US/09673032
; GENERAL INFORMATION:
; APPLICANT: MORGAN, BRYAN P.
; APPLICANT: RUSHMERE, NEIL K.
; APPLICANT: HINCHLIFE, STEWART J.
; APPLICANT: VAN DEN BERG, CARMEN W.
; TITLE OF INVENTION: MODIFIED BIOLOGICAL MATERIAL
; FILE REFERENCE: WN/KH/JT/MCM
; CURRENT FILING DATE: 2000-12-06
; PRIOR FILING DATE: 1999-04-08
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 128
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-673-032-20
```

```
Query Match
Best Local Similarity 100.0%; Score 96; DB 6; Length 128;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FEHCNFDVTRLRNE 17
Db 67 FEHCNFDVTRLRNE 83
```

```
RESULT 10
US-09-614-150-5766
; Sequence 5766, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Ventef, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-12-28
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5766
; LENGTH: 739
; TYPE: PRF
; ORGANISM: DROSOPHILA
US-09-614-150-5766
```

```
Query Match
Best Local Similarity 44.8%; Score 43; DB 6; Length 739;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 EHCNFDVTRLRN 16
Db 374 KHCNNTATIRKN 388
```

```
RESULT 11
US-09-991-630-12
; Sequence 12, Application US/09991630
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS T
; FILE REFERENCE: 540579-2007.3
; CURRENT APPLICATION NUMBER: US/09/991,630
; CURRENT FILING DATE: 2001-11-06
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-03-08
; PRIOR FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 236
; TYPE: PRF
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-991-630-12
```

```
Query Match
Best Local Similarity 43.2%; Score 41.5; DB 5; Length 236;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;
OY 3 HCNFDVTRLRNE 17
Db 42 HCNFDVTRLRNE 75
```

```
RESULT 12
US-09-991-630-15
; Sequence 15, Application US/09991630
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS TH
; FILE REFERENCE: 540579-2007.3
; CURRENT APPLICATION NUMBER: US/09/991,630
; CURRENT FILING DATE: 2001-11-06
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-03-08
; PRIOR FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 236
; TYPE: PRF
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-991-630-15
```

```
Query Match 43.2%; Score 41.5; DB 5; Length 236;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFNDVT-----TRLRENE 17
|||:|:|
DB 42 HCTFRDLTISIPDGPAVERVNLGYNLSLTRLTEND 75

RESULT 13
US-09-905-129-12
; Sequence 12, Application US/09905129
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
; FILE REFERENCE: 540579-2007.2
; CURRENT APPLICATION NUMBER: US/09/905,129
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/802,318
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/207,821
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/084,944
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/085,673
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-905-129-12

Query Match 43.2%; Score 41.5; DB 6; Length 236;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFNDVT-----TRLRENE 17
|||:|:|
DB 42 HCTFRDLTISIPDGPAVERVNLGYNLSLTRLTEND 75

RESULT 14
US-09-905-129-15
; Sequence 15, Application US/09905129
; GENERAL INFORMATION:
; APPLICANT: Elnat, et al
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
; FILE REFERENCE: 540579-2007.2
; CURRENT APPLICATION NUMBER: US/09/905,129
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/802,318
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/207,821
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/084,944
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/085,673
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236)
; OTHER INFORMATION: 'x' can be any amino acid
US-09-905-129-15

Query Match 43.2%; Score 41.5; DB 6; Length 236;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 3 HCNFNDVT-----TRLRENE 17
|||:|:|
DB 42 HCTFRDLTISIPDGPAVERVNLGYNLSLTRLTEND 75

RESULT 15
US-09-708-427-16838
; Sequence 16838, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16838
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..220
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..220
; OTHER INFORMATION: Ceres Seq. ID 1831334
US-09-708-427-16838

Query Match 41.7%; Score 40; DB 6; Length 220;
Best Local Similarity 46.2%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FEHCNFDVTTTL 13
|:||||:|
DB 60 FKHCNFSFVROL 72

RESULT 16
US-09-897-516-6395
; Sequence 6395, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kuesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6395
; LENGTH: 240
; TYPE: PRT
```

ORGANISM: Xenorhabdus sp.  
US-09-897-516-6395

## Query Match

Best Local Similarity 41.7%; Score 40; DB 6; Length 240;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 HCNENDVTTRLRENE 17  
| | | | | : |  
Db 62 HDTFNDVLRNLPHE 76

## RESULT 17

US-09-708-427-16837  
; Sequence 16837, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16837  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..244  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..244  
; OTHER INFORMATION: Ceres Seq. ID 1631333  
US-09-708-427-16837

## Query Match

Best Local Similarity 41.7%; Score 40; DB 6; Length 244;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRL 13  
| | | | | : |  
Db 84 FKHCNFSFVROL 96

## RESULT 18

US-60-336-049-36  
; Sequence 36, Application US/60336049  
; GENERAL INFORMATION:  
; APPLICANT: Adam, Luc  
; APPLICANT: Creelman, Robert  
; APPLICANT: Dubell, Arnold  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Keedle, James  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Reuber, Lynne  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: Genes for Modifying Plant Traits VI  
; FILE REFERENCE: MBI-0037  
; CURRENT APPLICATION NUMBER: US/60/336,049  
; CURRENT FILING DATE: 2001-12-05  
; NUMBER OF SEQ ID NOS: 784  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana

US-60-336-049-36

## Query Match

Best Local Similarity 41.7%; Score 40; DB 8; Length 244;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRL 13  
| | | | | : |  
Db 84 FKHCNFSFVROL 96

## RESULT 19

US-10-092-302-711  
; Sequence 711, Application US/10092302  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT218C1  
; CURRENT APPLICATION NUMBER: US/10/092,302  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 1040  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 711  
; LENGTH: 262  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-302-711

## Query Match

Best Local Similarity 41.7%; Score 40; DB 7; Length 262;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRL 11  
| | | | | : |  
Db 51 FEHCYFEDVTS 61

## RESULT 20

US-10-092-302-951  
; Sequence 951, Application US/10092302  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT218C1  
; CURRENT APPLICATION NUMBER: US/10/092,302  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 1040  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 951  
; LENGTH: 262  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-302-951

## Query Match

Best Local Similarity 41.7%; Score 40; DB 7; Length 262;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNENDVTTRL 11  
| | | | | : |  
Db 51 FEHCYFEDVTS 61

## RESULT 21

US-09-708-427-56000  
; Sequence 56000, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.

```

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56000
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..272
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..272
; OTHER INFORMATION: Ceres Seq. ID 1936468
US-09-708-427-56000
```

```

Query Match          41.7%; Score 40; DB 6; Length 272;
Best Local Similarity 46.2%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 FEHCNPNVDVTRL 13
    1:||||:         :1
Db 37 FKHCNFSFVROL 49
```

```

RESULT 22
US-09-708-427-55999
; Sequence 55999, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55999
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..299
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..299
; OTHER INFORMATION: Ceres Seq. ID 1936467
US-09-708-427-55999
```

```

Query Match          41.7%; Score 40; DB 6; Length 299;
Best Local Similarity 46.2%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 FEHCNPNVDVTRL 13
    1:||||:         :1
Db 64 FKHCNFSFVROL 76
```

```

RESULT 23
PCT-US02-03987-15325
; Sequence 15325, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028VPC
```

```

; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15325
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Candida albicans
PCT-US02-03987-15325
```

```

Query Match          41.7%; Score 40; DB 1; Length 394;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy 5 NPNVDVTRLRENE 17
    |||:|:  |||:|
Db 349 NFNNVSAELKERE 361
```

```

RESULT 24
US-10-032-585-7381
; Sequence 7381, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7381
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7381
```

```

Query Match          41.7%; Score 40; DB 7; Length 394;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy 5 NPNVDVTRLRENE 17
    |||:|:  |||:|
Db 349 NFNNVSAELKERE 361
```

```

RESULT 25
US-10-072-851-15325
; Sequence 15325, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibit
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
```

```

; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15325
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-072-851-15325

```

```

Query Match      41.7%; Score 40; DB 7; Length 394;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 5 NNDVTTRLRENE 17
   |||:|: |:| |
Db 349 NNNVSALERE 361

```

Search completed: April 2, 2002, 09:23:05  
 Job time: 194 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:19:51 ; Search time 12.43 Seconds  
(without alignments)  
30.777 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRLRENE 17

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 96    | 100.0       | 103    | 1     | US-08-271-562-1   |
| 2          | 96    | 100.0       | 103    | 1     | US-08-087-007-3   |
| 3          | 96    | 100.0       | 103    | 2     | US-08-696-777-1   |
| 4          | 96    | 100.0       | 103    | 3     | US-08-483-433-3   |
| 5          | 96    | 100.0       | 103    | 5     | PCT-US92-05920-3  |
| 6          | 96    | 100.0       | 105    | 4     | US-09-591-435-12  |
| 7          | 96    | 100.0       | 128    | 6     | 5179198-1         |
| 8          | 96    | 100.0       | 128    | 6     | 5521296-1         |
| 9          | 81    | 84.4        | 121    | 4     | US-09-591-435-13  |
| 10         | 81    | 84.4        | 1162   | 2     | US-08-663-566A-15 |
| 11         | 41    | 42.7        | 1162   | 2     | US-08-023-610-15  |
| 12         | 41    | 42.7        | 1162   | 2     | US-08-288-065A-15 |
| 13         | 41    | 42.7        | 1162   | 2     | US-08-362-240A-15 |
| 14         | 41    | 42.7        | 1162   | 2     | PCT-US95-10245-15 |
| 15         | 41    | 42.7        | 1162   | 2     | US-08-277-231A-12 |
| 16         | 41    | 41.7        | 217    | 1     | US-08-473-750-1   |
| 17         | 40    | 41.7        | 217    | 2     | US-08-473-750-5   |
| 18         | 40    | 41.7        | 217    | 2     | US-08-473-750-5   |
| 19         | 40    | 41.7        | 217    | 2     | US-08-477-326-1   |
| 20         | 40    | 41.7        | 217    | 2     | US-08-477-326-5   |
| 21         | 39.5  | 41.1        | 602    | 4     | US-09-257-490-15  |
| 22         | 38    | 39.6        | 278    | 1     | US-08-392-828C-4  |
| 23         | 37    | 38.5        | 278    | 3     | US-09-330-945-4   |
| 24         | 37    | 38.5        | 586    | 2     | US-08-630-822A-70 |
| 25         | 36.5  | 38.0        | 486    | 2     | US-09-005-069-70  |
| 26         | 36.5  | 38.0        | 486    | 3     | US-08-889-841B-8  |
| 27         | 36.5  | 38.0        | 491    | 3     | US-08-889-841B-10 |

|    |      |      |      |   |                   |                   |
|----|------|------|------|---|-------------------|-------------------|
| 28 | 36   | 37.5 | 206  | 4 | US-09-156-580-2   | Sequence 2, App1  |
| 29 | 36   | 37.5 | 298  | 1 | US-08-393-985-10  | Sequence 10, App1 |
| 30 | 36   | 37.5 | 300  | 4 | US-09-433-598-2   | Sequence 2, App1  |
| 31 | 36   | 37.5 | 312  | 4 | US-08-821-872-2   | Sequence 2, App1  |
| 32 | 36   | 37.5 | 364  | 1 | US-08-680-726A-56 | Sequence 56, App1 |
| 33 | 36   | 37.5 | 364  | 4 | US-09-092-409-56  | Sequence 56, App1 |
| 34 | 36   | 37.5 | 421  | 2 | US-08-576-626A-53 | Sequence 53, App1 |
| 35 | 36   | 37.5 | 469  | 3 | US-08-985-335-5   | Sequence 5, App1  |
| 36 | 36   | 37.5 | 469  | 3 | US-08-985-335-9   | Sequence 5, App1  |
| 37 | 36   | 37.5 | 469  | 4 | US-09-410-372-5   | Sequence 9, App1  |
| 38 | 36   | 37.5 | 469  | 4 | US-09-410-372-9   | Sequence 9, App1  |
| 39 | 36   | 37.5 | 629  | 4 | US-09-300-909-19  | Sequence 31, App1 |
| 40 | 36   | 37.5 | 1114 | 2 | US-08-576-626A-31 | Sequence 31, App1 |
| 41 | 36   | 37.5 | 1233 | 4 | US-09-194-613-5   | Sequence 2, App1  |
| 42 | 36   | 37.5 | 1694 | 1 | US-08-494-168-2   | Sequence 2, App1  |
| 43 | 35.5 | 37.0 | 455  | 3 | US-08-889-841B-46 | Sequence 46, App1 |
| 44 | 35.5 | 37.0 | 474  | 3 | US-08-889-841B-36 | Sequence 36, App1 |
| 45 | 35.5 | 37.0 | 474  | 3 | US-08-889-841B-39 | Sequence 39, App1 |
| 46 | 35.5 | 37.0 | 536  | 2 | US-08-392-806A-2  | Sequence 2, App1  |
| 47 | 35.5 | 37.0 | 536  | 4 | US-09-257-490-2   | Sequence 2, App1  |
| 48 | 35.5 | 37.0 | 600  | 2 | US-08-392-806A-4  | Sequence 4, App1  |
| 49 | 35.5 | 37.0 | 600  | 4 | US-09-257-490-4   | Sequence 4, App1  |
| 50 | 35.5 | 37.0 | 605  | 2 | US-08-472-666-1   | Sequence 1, App1  |

#### ALIGNMENTS

RESULT 1  
US-08-271-562-1  
; Sequence 1, Application US/08271562  
; Patent No. 5573940  
; GENERAL INFORMATION:  
; APPLICANT: Sims, Peter J.  
; APPLICANT: Bothwell, Alfred L.M.  
; TITLE OF INVENTION: Genetic Inhibition of Complement  
; TITLE OF INVENTION: Mediated Inflammatory Response  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: U.S.  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/271,562  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/729926  
; FILING DATE: 15-JUL-1991  
; APPLICATION NUMBER: US 07/365199  
; FILING DATE: 04-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: OMRF 112CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 815-6508  
; TELEFAX: (404) 815-6555  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 103 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

;;  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; TISSUE TYPE: Blood  
;; CELL TYPE: Erythrocyte  
US-08-271-562-1

Query Match 100.0%; Score 96; DB 1; Length 103;  
Best Local Similarity 100.0%; Pred. No. 4.8e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTRLRENE 17  
DB 42 FEHCNFNDVTTRLRENE 58

## RESULT 2

US-08-087-007-3  
; Sequence 3, Application US/08087007  
; Patent No. 5705732  
; Patent No. 5705732 5684223  
; GENERAL INFORMATION:

APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
APPLICANT: Elliott, Eileen A.  
APPLICANT: Flavell, Richard A.  
APPLICANT: Madril, Joseph  
APPLICANT: Rollins, Scott  
APPLICANT: Squinto, Stephen  
TITLE OF INVENTION: Universal Donor Cells  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.

ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 19930701  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: CD59  
US-08-087-007-3

Query Match 100.0%; Score 96; DB 1; Length 103;  
Best Local Similarity 100.0%; Pred. No. 4.8e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTRLRENE 17  
DB 42 FEHCNFNDVTTRLRENE 58

## RESULT 3

US-08-696-777-1  
; Sequence 1, Application US/08696777  
; Patent No. 5955441  
; GENERAL INFORMATION:

APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
TITLE OF INVENTION: Genetic Inhibition of Complement Mediated  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2600 One Atlantic Center, 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 13-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR 112cipdlv  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-873-8794  
TELEFAX: (404)-873-8795  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-696-777-1

Query Match 100.0%; Score 96; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 4.8e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFNDVTTRLRENE 17  
DB 42 FEHCNFNDVTTRLRENE 58

## RESULT 4

US-08-483-433-3  
; Sequence 3, Application US/08483433  
; Patent No. 6100443  
; GENERAL INFORMATION:

APPLICANT: Sims et al.  
TITLE OF INVENTION: Universal Donor Cells  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,433  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/087,007  
FILING DATE: July 1, 1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/906,394  
FILING DATE: June 29, 1992  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF135c1p2 div  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-9794  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: CD59  
US-08-483-433-3

Query Match 100.0%; Score 96; DB 3; Length 103;  
Best Local Similarity 100.0%; Pred. No. 4,8e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRLENE 17  
|||||

Db 42 FEHCNFDVTRLENE 58

RESULT 5  
PCT-US92-05920-3  
Sequence 3, Application PC/TUS9205920  
GENERAL INFORMATION:  
APPLICANT: Sims, Peter J.  
APPLICANT: Bothwell, Alfred L.M.  
APPLICANT: Elliott, Eileen A.  
APPLICANT: Flavell, Richard A.  
APPLICANT: Maertl, Joseph  
APPLICANT: Rollins, Scott  
APPLICANT: Bell, Leonard  
APPLICANT: Squinto, Stephen  
TITLE OF INVENTION: Universal Donor Cells  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05920  
FILING DATE: 19920714  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6500  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 103 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: CD59  
PCT-US92-05920-3

Query Match 100.0%; Score 96; DB 5; Length 103;  
Best Local Similarity 100.0%; Pred. No. 4,8e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FEHCNFDVTRLENE 17  
|||||

Db 42 FEHCNFDVTRLENE 58

RESULT 6  
US-09-591-435-12  
Sequence 12, Application US/09591435  
Patent No. 6280953  
GENERAL INFORMATION:  
APPLICANT: MESSIER, WALTER  
APPLICANT: SIKELA, JAMES M  
TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE  
SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL  
TITLE OF INVENTION: AND MEDICAL CONDITIONS  
FILE REFERENCE: GENO.200.2  
CURRENT APPLICATION NUMBER: US/09/591,435  
CURRENT FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 09/591,435  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 09/240,915  
PRIOR FILING DATE: 1999-01-29  
PRIOR APPLICATION NUMBER: 60/073,263  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/098,987  
PRIOR FILING DATE: 1998-09-02  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 12  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Homo sapiens

US-09-591-435-12

Query Match  
Best Local Similarity 100.0%; Score 96; DB 4; Length 105;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNNDVTRLRNE 17  
Db 43 FEHCNNDVTRLRNE 59

RESULT 7  
5179198-1  
; Patent No. 5179198  
; APPLICANT: OKADA, HIDECHIKA; OKADA, NORIKO; NAGAMI, YOICHI;  
; TAKASHI, KAZUHIRO; TAKIZAWA, HISAO; KONDO, JUN  
; TITLE OF INVENTION: GLYCOPROTEIN AND GENE CODING THEREFOR  
; NUMBER OF SEQUENCES: 17  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/376,828  
; FILING DATE: 07-JUL-1989  
; SEQ ID NO:1:  
; LENGTH: 128  
5179198-1

Query Match  
Best Local Similarity 100.0%; Score 96; DB 6; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNNDVTRLRNE 17  
Db 67 FEHCNNDVTRLRNE 83

RESULT 8  
5521296-1  
; Patent No. 5521296  
; APPLICANT: OKADA, HIDECHIKA; OKADA, NORIKO; NAGAMI, YOICHI;  
; TAKAHASHI, KAZUHIRO; TAKIZAWA, HISAO; KONDO, JUN  
; TITLE OF INVENTION: GLYCOPROTEIN AND GENE CODING THEREFOR  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/739,211  
; FILING DATE: 01-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 376,828  
; FILING DATE: 07-JUL-1989  
; SEQ ID NO:1:  
; LENGTH: 128  
5521296-1

Query Match  
Best Local Similarity 100.0%; Score 96; DB 6; Length 128;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHCNNDVTRLRNE 17  
Db 67 FEHCNNDVTRLRNE 83

RESULT 9  
US-09-591-435-13  
; Sequence 13; Application US/09591435  
; Patent No. 6280953  
; GENERAL INFORMATION:  
; APPLICANT: MESSIER, WALTER  
; APPLICANT: SIKELA, JAMES M  
; TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL  
; TITLE OF INVENTION: AND MEDICAL CONDITIONS

FILE REFERENCE: GENO.200.2  
; CURRENT APPLICATION NUMBER: US/09/591,435  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 09/591,435  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 09/240,915  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: 60/073,263  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/098,987  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 121  
; TYPE: PRT  
; ORGANISM: Pan troglodytes  
US-09-591-435-13

Query Match  
Best Local Similarity 84.4%; Score 81; DB 4; Length 121;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 EHCNFKDITRLRNE 17  
Db 44 EHCNFKDITRLRNE 59

RESULT 10  
US-08-663-566A-15  
; Sequence 15; Application US/08663566A  
; Patent No. 5853733  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: MacDonald, Richard D  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/663,566A  
; FILING DATE: June 13, 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)278-0400  
; TELEFAX: (212)391-0526  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-663-566A-15

Query Match  
Best Local Similarity 42.7%; Score 41; DB 2; Length 1162;  
Matches 75.0%; Pred. No. 97;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
|||||  
Db 101 HCNFSDDT 108

RESULT 11  
US-08-023-610-15  
Sequence 15, Application US/08023610  
Patent No. 5928648  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Macdonald Ph.D., Richard D  
TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys  
TITLE OF INVENTION: and Uses Thereof  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/023,610  
FILING DATE: February 26, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White Esq, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-023-610-15

Query Match 42.7%; Score 41; DB 2; Length 1162;  
Best Local Similarity 75.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
|||||  
Db 101 HCNFSDDT 108

RESULT 12  
US-08-288-065A-15  
Sequence 15, Application US/0828065A  
Patent No. 5961362  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D  
APPLICANT: Macdonald, Richard D  
TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys S-  
TITLE OF INVENTION: HVT-050 and Uses Thereof  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/288,065A  
FILING DATE: Aug-09-94  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-288-065A-15

Query Match 42.7%; Score 41; DB 2; Length 1162;  
Best Local Similarity 75.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
|||||  
Db 101 HCNFSDDT 108

RESULT 13  
US-08-362-240A-15  
Sequence 15, Application US/08362240A  
Patent No. 5965138  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D  
APPLICANT: Junker, David  
TITLE OF INVENTION: Recombinant Herpesvirus and Uses Thereof  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,240A  
FILING DATE: Dec-22-94  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-362-240A-15

Query Match  
Best Local Similarity 42.7%; Score 41; DB 2; Length 1162;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HCNFNDVT 10  
DB 101 HCNFSDTT 108

## RESULT 14

PCT-US95-10245-15  
Sequence 15, Application PC/TUS9510245  
GENERAL INFORMATION:  
APPLICANT: SYNTRO CORPORATION  
TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10245  
FILING DATE: 09-AUG-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-10245-15

Query Match  
Best Local Similarity 42.7%; Score 41; DB 5; Length 1162;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 HCNFNDVT 10  
DB 101 HCNFSDTT 108

## RESULT 15

US-08-277-231A-2  
Sequence 2, Application US/08277231A  
Patent No. 5643725  
GENERAL INFORMATION:  
APPLICANT: Green, Bruce A.  
TITLE OF INVENTION: Sequence and Analysis of LKP Pili  
Patent No. 5643725  
TITLE OF INVENTION: Structural Genes and The LKP Pili Operon of No. 5643725typable  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/277,231A  
FILING DATE: 19-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 32,542  
REFERENCE/DOCKET NUMBER: ACC94-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-277-231A-2

Query Match  
Best Local Similarity 41.7%; Score 40; DB 1; Length 217;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
QY 1 FHCNNDVTYRLRENE 17  
DB 156 FTHSNNDVATQOTVVK 172

## RESULT 16

US-08-277-231A-12  
Sequence 12, Application US/08277231A  
Patent No. 5643725  
GENERAL INFORMATION:  
APPLICANT: Green, Bruce A.  
TITLE OF INVENTION: Sequence and Analysis of LKP Pili  
Patent No. 5643725  
TITLE OF INVENTION: Structural Genes and The LKP Pili Operon of No. 5643725typable  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/277,231A  
FILING DATE: 19-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 32,542  
REFERENCE/DOCKET NUMBER: ACC94-02  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-277-231A-12

Query Match 41.7%; Score 40; DB 1; Length 217;  
Best Local Similarity 47.1%; Pred. No. 22;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 FEHCNENDVTRLRNE 17  
| | | | | : 1:  
DB 156 FTHSNTNDVATQOTVVK 172

RESULT 17  
US-08-473-750-1  
; Sequence 1, Application US/08473750  
; Patent No. 5834187  
; Patent No. 5834187 5786143  
; GENERAL INFORMATION:  
; APPLICANT: Green, Bruce A.  
; APPLICANT: Brinton, Jr., Charles C.  
; TITLE OF INVENTION: Sequence and Analysis of LKP Pilln  
; Patent No. 5834187  
; TITLE OF INVENTION: Structural Gene and the LKP Pill Operon of No. 5834187 5786143  
; TITLE OF INVENTION: Haemophilus Influenzae  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,750  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/277,321  
; FILING DATE: 19-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Alice O.  
; REGISTRATION NUMBER: 33,542  
; REFERENCE/DOCKET NUMBER: ACC94-02B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-473-750-1

Query Match 41.7%; Score 40; DB 2; Length 217;  
Best Local Similarity 47.1%; Pred. No. 22;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 FEHCNENDVTRLRNE 17  
| | | | | : 1:  
DB 156 FTHSNTNDVATQOTVVK 172

RESULT 18  
US-08-473-750-5  
; Sequence 5, Application US/08473750  
; Patent No. 5834187  
; Patent No. 5834187 5786143  
; GENERAL INFORMATION:  
; APPLICANT: Green, Bruce A.  
; APPLICANT: Brinton, Jr., Charles C.  
; TITLE OF INVENTION: Sequence and Analysis of LKP Pilln  
; Patent No. 5834187  
; TITLE OF INVENTION: Structural Gene and the LKP Pill Operon of No. 5834187 5786143  
; TITLE OF INVENTION: Haemophilus Influenzae  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,750  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/277,321  
; FILING DATE: 19-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Alice O.  
; REGISTRATION NUMBER: 33,542  
; REFERENCE/DOCKET NUMBER: ACC94-02B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-473-750-5

Query Match 41.7%; Score 40; DB 2; Length 217;  
Best Local Similarity 47.1%; Pred. No. 22;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 FEHCNENDVTRLRNE 17  
| | | | | : 1:  
DB 156 FTHSNTNDVATQOTVVK 172

RESULT 19  
US-08-477-326-1  
; Sequence 1, Application US/08477326  
; Patent No. 5968769  
; GENERAL INFORMATION:  
; APPLICANT: Green, Bruce A.  
; APPLICANT: Brinton, Jr., Charles C.  
; TITLE OF INVENTION: Sequence and Analysis of LKP Pilln  
; Patent No. 5968769

```

; TITLE OF INVENTION: Structural Gene and the LKP Pili Operon of No. 5968769lypable
; NUMBER OF INVENTION: Haemophilus Influenzae
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,326
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/277,231
; FILING DATE: July 19, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: ACC94-02A
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-477-326-1

Query Match
Best Local Similarity 41.7%; Score 40; DB 2; Length 217;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FEHCFNDVTRLRNE 17
Db 156 FTSHNDVATQOTVKN 172

RESULT 20
; US-08-477-326-5
; Sequence 5, Application US/08477326
; Patent No. 5968769
; GENERAL INFORMATION:
; APPLICANT: Green, Bruce A.
; APPLICANT: Brinton, Jr., Charles C.
; TITLE OF INVENTION: Sequence and Analysis of LKP Pili
; Patent No. 5968769
; TITLE OF INVENTION: Structural Gene and the LKP Pili Operon of No. 5968769lypable
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/477,326
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/277,231
; FILING DATE: July 19, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: ACC94-02A
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-326-5
```

```

Query Match
Best Local Similarity 41.7%; Score 40; DB 2; Length 217;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FEHCFNDVTRLRNE 17
Db 156 FTSHNDVATQOTVKN 172
```

```

RESULT 21
; US-09-257-490-15
; Sequence 15, Application US/09257490A
; Patent No. 6248328
; GENERAL INFORMATION:
; APPLICANT: Dietrich, Ursula
; APPLICANT: Von Briesen, Hagen
; APPLICANT: Grez, Manuel
; APPLICANT: Rubsamen-Maigmann, Helga
; TITLE OF INVENTION: HIV-1 virus isolates of a subtype and its differential
; TITLE OF INVENTION: of this subtype, a vaccine against HIV-1 virus infections
; TITLE OF INVENTION: of this subtype and method of producing same, use of
; FILE REFERENCE: 10496/P58512051
; CURRENT APPLICATION NUMBER: US/09/257,490A
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 602
; TYPE: PRF
; ORGANSIM: Simian Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: SIVCPZ
; US-09-257-490-15
```

```

Query Match
Best Local Similarity 41.1%; Score 39.5; DB 4; Length 602;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 2 EHCNFDVTRLRNE 17
Db 45 KNCSPN-VTTELDRKK 59
```

```

RESULT 22
; US-08-392-828C-4
; Sequence 4, Application US/08392828C
; Patent No. 5795962
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SADAOKI
```



APPLICANT: MOTA, TATSUSHI  
APPLICANT: SEKI, NORIAKI  
APPLICANT: ODA, TOSHIO  
TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
ADDRESSEE: THIBEAULT  
STREET: 53 STATE STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,828C  
FILING DATE: 28-FEB-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: FJN-033  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-392-828C-4

Query Match 39.6%; Score 38; DB 1; Length 278;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EHCNFDVTTRL 13  
Db 5 KHCGRPVITRI 16

RESULT 23  
US-09-330-945-4  
Sequence 4, Application US/09330945  
Patent No. 6077946  
GENERAL INFORMATION:  
APPLICANT: IWANAGA, SADAKI  
APPLICANT: MOTA, TATSUSHI  
APPLICANT: SEKI, NORIAKI  
APPLICANT: ODA, TOSHIO  
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB  
TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
ADDRESSEE: THIBEAULT, LLP  
STREET: 125 HIGH STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/330,945  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION NUMBER: 09/119,995  
APPLICATION NUMBER: 09/119,995  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER, EDMUND R  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: FJN-032DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-330-945-4

Query Match 39.6%; Score 38; DB 3; Length 278;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EHCNFDVTTRL 13  
Db 5 KHCGRPVITRI 16

RESULT 24  
US-08-630-822A-70  
Sequence 70, Application US/08630822A  
Patent No. 5840695  
GENERAL INFORMATION:  
APPLICANT: FRANK, GLENN R.  
APPLICANT: HUNTER, SHIRLEY WU  
APPLICANT: WALLENFELS, LYNDIA  
TITLE OF INVENTION: NOVEL ECTOPARASITE SALLIYA PROTEINS  
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,822A  
FILING DATE: 11-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CONNELL, GARY J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-17-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 586 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:

Tue Apr 2 09:51:55 2002

us-09-020-393b-3\_copy\_42\_58.open.ra1

NAME/KEY: Xaa = any amino acid  
LOCATION: 379  
US-08-630-822A-70

Search completed: April 2, 2002, 09:20:32  
Job time: 41 sec

Query Match  
Best Local Similarity 38.5%; Score 37; DB 2; Length 586;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 EHCNENDVT 10  
Db 578 EHCYDPDT 586

RESULT 25

US-09-005-069-70  
; Sequence 70, Application US/09005069  
; Patent No. 5932470  
; GENERAL INFORMATION:  
; APPLICANT: FRANK, GLENN R.  
; APPLICANT: HUNTER, SHIRLEY WU  
; APPLICANT: WALLBENELS, LYNDIA  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
; NUMBER OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,069  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/630,822  
; FILING DATE: 11-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CONNELL, GARY J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 586 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Xaa = any amino acid  
; LOCATION: 379  
US-09-005-069-70

Query Match  
Best Local Similarity 38.5%; Score 37; DB 2; Length 586;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 2 EHCNENDVT 10  
Db 578 EHCYDPDT 586

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:01:52 ; Search time 37.51 Seconds  
(without alignments)  
54.831 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145

Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description          |
|------------|-------|-------------|--------|----------|----------------------|
| 1          | 142   | 97.9        | 559    | 1 C9HU   | complement C9 prec   |
| 2          | 52    | 35.9        | 3237   | 2 T37964 | probable ubiquitin   |
| 3          | 49    | 33.8        | 135    | 2 PN0494 | NAD+ ADP-ribosyltr   |
| 4          | 49    | 33.8        | 524    | 2 A29677 | complement C9 prec   |
| 5          | 46.5  | 32.1        | 1002   | 2 T19226 | hypothetical prote   |
| 6          | 46    | 31.7        | 203    | 2 T30493 | hypothetical prote   |
| 7          | 46    | 31.7        | 254    | 2 E84826 | hypothetical prote   |
| 8          | 46    | 31.7        | 292    | 2 C70421 | conserved hypotet    |
| 9          | 46    | 31.7        | 357    | 2 S27909 | hypothetical prote   |
| 10         | 46    | 31.7        | 357    | 2 JQ2174 | hypothetical prote   |
| 11         | 46    | 31.7        | 930    | 2 JX0368 | inter-alpha-trypsi   |
| 12         | 46    | 31.7        | 1372   | 2 T25933 | hypothetical prote   |
| 13         | 46    | 31.7        | 2144   | 2 T21712 | hypothetical prote   |
| 14         | 45.5  | 31.4        | 661    | 2 C81822 | topoisomerase IV s   |
| 15         | 45.5  | 31.4        | 661    | 2 D81055 | topoisomerase IV c   |
| 16         | 45    | 31.0        | 358    | 2 T44333 | hypothetical prote   |
| 17         | 45    | 31.0        | 440    | 2 C84265 | adenylosuccinate s   |
| 18         | 45    | 31.0        | 732    | 2 T18567 | hypothetical prote   |
| 19         | 45    | 31.0        | 895    | 2 T49010 | hypothetical prote   |
| 20         | 45    | 31.0        | 1189   | 2 T30319 | hian-Aal retrorran   |
| 21         | 44.5  | 30.7        | 884    | 2 T18649 | hypothetical prote   |
| 22         | 44    | 30.3        | 241    | 2 T26909 | hypothetical prote   |
| 23         | 44    | 30.3        | 261    | 2 A86763 | cobyrinic acid synth |
| 24         | 44    | 30.3        | 535    | 2 F81847 | dihydrolipeamide S   |
| 25         | 44    | 30.3        | 535    | 2 F81094 | pyruvate dehydroge   |
| 26         | 44    | 30.3        | 560    | 2 S42158 | KREI1 protein - ye   |
| 27         | 44    | 30.3        | 659    | 2 A81862 | transketolase (EC    |
| 28         | 44    | 30.3        | 659    | 2 B81082 | transketolase NMB1   |
| 29         | 44    | 30.3        | 668    | 2 A46013 | coagulation factor   |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 44   | 30.3 | 678  | 2 A71287 | probable cytoplasm |
| 31 | 44   | 30.3 | 679  | 2 S54299 | transketolase (EC  |
| 32 | 44   | 30.3 | 944  | 1 S48821 | probable membrane  |
| 33 | 44   | 30.3 | 971  | 2 E96794 | hypothetical prote |
| 34 | 43.5 | 30.0 | 117  | 2 H75165 | hypothetical prote |
| 35 | 43.5 | 30.0 | 247  | 2 T52511 | related to lysopho |
| 36 | 43.5 | 30.0 | 341  | 2 S64618 | hypothetical prote |
| 37 | 43.5 | 30.0 | 739  | 2 F86337 | hypothetical prote |
| 38 | 43.5 | 30.0 | 884  | 2 E86244 | unknown protein, 4 |
| 39 | 43.5 | 30.0 | 2550 | 2 B53435 | vesicular transpor |
| 40 | 43   | 29.7 | 201  | 2 A47096 | response regulator |
| 41 | 43   | 29.7 | 257  | 2 C64442 | hypothetical prote |
| 42 | 43   | 29.7 | 305  | 2 F64481 | hypothetical prote |
| 43 | 43   | 29.7 | 309  | 2 T09564 | glutaminyl-peptide |
| 44 | 43   | 29.7 | 311  | 2 S52498 | GPW2 protein - yea |
| 45 | 43   | 29.7 | 323  | 2 T26536 | hypothetical prote |

## ALIGNMENTS

RESULT 1  
C9HU  
complement C9 precursor [validated] - human  
C:Species: Homo sapiens (man)  
C:Date: 27-Nov-1985 #sequence.revision 17-Nov-2000 #text.change 17-Nov-2000  
C:Accession: A59363; I52400; A91029; A94019; S68647; A59364; A03208  
R:Maraziti, D.; Eggertsen, G.; Fey, G.H.; Stanley, K.K.  
unpublished results, 1988, cited by GenBank  
A:Title: Relationships between the gene and protein structure in human compleme  
A:Reference number: A59363  
A:Accession: A59363  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-559 <MAR1>  
A:Cross-references: GB:X02176; NID:929580; PIDN:CAA26117.1; PID:929581  
R:Maraziti, D.; Eggertsen, G.; Fey, G.H.; Stanley, K.K.  
Biochemistry 27, 6529-6534, 1988  
A:Title: Relationships between the gene and protein structure in human complem  
A:Reference number: I52400; MUID:89118250  
A:Accession: I52400  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 62-159 <MAR2>  
A:Cross-references: GB:X02833; NID:9179727; PIDN:AAA51890.1; PID:9179728  
R:Stanley, K.K.; Kocher, H.P.; Luzzio, J.P.; Jackson, P.; Tschopp, J.  
EMBO J. 4, 375-382, 1985  
A:Title: The sequence and topology of human complement component C9.  
A:Reference number: A91029; MUID:85257464  
A:Accession: A91029  
A:Molecule type: mRNA  
A:Residues: 5',1-313,315-559 <STA>  
A:Cross-references: GB:X02176; NID:929580  
R:Disclipio, R.G.; Gehring, M.R.; Podack, E.R.; Kan, C.C.; Hugli, T.E.; Fey, G.H.  
Proc. Natl. Acad. Sci. U.S.A. 81, 7299-7302, 1984  
A:Title: Nucleotide sequence of cDNA and derived amino acid sequence of human complem  
A:Reference number: A94019; MUID:85063778  
A:Accession: A94019  
A:Molecule type: mRNA  
A:Residues: 2-12,'X',14-16,'X',18-42,'R',44-313,315-416,'P',418-559 <DIS>  
A:Cross-references: GB:X02766; NID:9179725; PIDN:AAA51889.1; PID:9179726  
R:Langweiler, S.; Schaller, J.; Rickli, E.E.  
FEBS Lett. 380, 8-12, 1996  
A:Title: Identification of disulfide bonds in the ninth component (C9) of human compl  
A:Reference number: S68647; MUID:96181657  
A:Accession: S68647  
A:Molecule type: protein  
A:Residues: 34-47,'S',52-59,'G',69-87,'X',89-93,'T',94-98,'106-113,118-131,136-145,180-181,'X'  
R:Witze, Schlomp, K.; Hobart, M.J.; Fernie, B.A.; Orren, A.; Wurzner, R.; Rittner, C.;  
Immunogenetics 48, 144-147, 1998  
A:Title: Heterogeneity in the genetic basis of human complement C9 deficiency.  
A:Reference number: A59364; MUID:98296010  
A:Note: submitted to GenBank, September 1996

A:Accession A59364  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 27-559 <MIT>  
A:Cross-references: GB:Y08545; NID:Y1634472; PIDN:CA69649.1; PID:q2258128  
R:Experimental source: Macromolecular tissue type blood; cell type lymphocyte  
J. Biol. Chem. 274 32786-32794, 1999  
A:Title: The four terminal components of the complement system are C-mannosylated on mu  
A:Reference number: A59362; MUID:20020247  
A:Contents: annotation  
A:Note: Identification and location of C-mannosylation sites by mass-spectroscopy and (1  
C:Genetics:  
A:Gene: GDB:G9  
A:Cross-references: GDB:119738; OMIM:120940  
A:Map position: 5p13-5p13  
C:Complex: monomer in plasma; 10 to 16 chain multimer in transmembrane form  
A:Description: In association with complement C5b-8 complex polymerizes to form a trans  
A:Pathway: complement pathway  
C:Superfamily: complement C9; EGF homology; LDL receptor ligand-binding repeat homology;  
C:Keywords: complement C9; EGF homology; LDL receptor ligand-binding repeat homology;  
F:1-21/Domain: signal sequence; cytolysis; glycoprotein; inflammation; membrane attack C  
F:22-265/Domain: product: complement C9 #status predicted <SIG>  
F:22-265/Domain: complement C9a #status predicted <CAT>  
F:101-134/Domain: thrombospondin type 1 repeat homology <THR>  
F:266-559/Domain: LDL receptor ligand-binding repeat homology <LDL>  
F:510-539/Domain: complement C9b #status predicted <C9B>  
F:43-78,54-57,88-94,142-181,254-255,380-405,510-526,513-528,530-539/Disulfide bonds: #st  
F:48/Modified site: 2' mannosyl-tryptophan (Trp) #status experimental  
F:51/Modified site: 2' mannosyl-tryptophan (Trp) (partial) #status experimental  
F:265-266/cleavage site: His-Gly (thrombin) #status predicted  
F:277,415/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 97.9% Score 142; DB 1; Length 559;  
Best Local Similarity 96.3%; Pred. No. 2,66-14;  
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 KCLGYHLDVSLAFSEISGAEFNKDDC 27  
Db 379 RCLGYHLDVSLAFSEISGAEFNKDDC 405  
:|||||  
A:Accession: T37964  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3227 <DEV>  
A:Cross-references: EMBL:Z99531; PIDN:CA16714.1; GSPDB:GN00066; SPDB:SPAC19D5.04  
A:Experimental source: Strain 972h-? cosmid c19D5  
C:Genetics:  
A:Gene: SPDB:SPAC19D5.04  
A:Map position: 1

```

RESULT      3
PN0494
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Cherry salmon (fragment)
N:Alternate names: poly ADP-ribose polymerase
C:Species: Oncorhynchus masou (cherry salmon)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #extl_change 17-Mar-1999
C:Accession: PN0494
R:Ozawa, Y.; Uchida, K.; Aml, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A:Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose) polymerase
A:Reference number: PN0494; MUID:93277538
A:Accession: PN0494
A:Molecule type: mRNA
A:Residues: 1-135 <COZA>
C:Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and dif
C:Superfamily: NAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pent

```

RESULT 4  
A29677  
complement C9 precursor - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 16-Jul-1999  
C:Accession: A29677  
R:Stanley, K.K.; Herz, J.  
EMBO J. 6, 1951-1957, 1987  
A:Title: Topological mapping of complement component C9 by recombinant DNA techniques  
A:Reference number: A29677; MUID:88004404  
A:Accession: A29677  
A:Molecule type: mRNA  
A:Residues: 1-524 <STA>  
A:Cross-references: GB:X05475  
C:Superfamily: complement C9; EGF homology; LDL receptor ligand-binding repeat homolo  
E:19-73/Domain: complement alternate pathway; complement pathway; cytolysis; glycoprotein  
F:79-112/Domain: LDL receptor ligand-binding repeat homology <THR>  
F:489-518/Domain: EGF homology <EGF>

| Query Match           | 33.8%                          | Score 49        | DB 2          | Length 524 |
|-----------------------|--------------------------------|-----------------|---------------|------------|
| Best Local Similarity | 33.3%                          | Pred. No. 11    |               |            |
| Matches 10            | Conservative                   | 6               | Mismatches 10 | Indels 4   |
|                       |                                |                 |               | Gaps 1     |
| QY 2                  | CLGHLDVSLA----                 | FSEISYCAEENKDDC | 27            |            |
|                       | : :                            | : : :           |               |            |
| Db 354                | CLGEMDLRIPIODDLKADSVTASVNAADGC | 383             |               |            |

RESULT 5  
 T19226  
 hypothetical protein C13B4.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision  
 C:Accession: T19226  
 R:SMY, R.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19093  
 A:Accession: T19226  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1002 <MIL>  
 A:Cross-references: EMBL:Z81468; PIDN:CAE03875.2; GSPDB:GN00020; CESP:C13B4.1.







|                       |                 |               |           |             |
|-----------------------|-----------------|---------------|-----------|-------------|
| Query Match           | 31.48;          | Score 45.5;   | DB 2;     | Length 661; |
| Best Local Similarity | 34.58;          | Pred. NO. 49; |           |             |
| Matches 10;           | Conservative 5; | Mismatches 9; | Indels 5; | Gaps 1.     |

```

QY      3 LGYHLD-----VSLAFSEISVGAEFNKDD 26
      :| | : | | : | :| | |
Db      78 VGLHPEEGVPVELVFTTRLHAGGKFNKKD 106

```

**RESULT 15**

topology: merge\_v chain B NM01682 [imported] - Neisseria meningitidis (strain MC58 serogroup C)  
 C: Species: Neisseria meningitidis  
 C: Date: 31-Mar-2000 #sequence: revision 31-Mar-2000 #text: change 19-Jan-2001  
 C: Accession: D81055  
 R: Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.R.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Arif, H.; Qiu, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M. Science 287, 1809-1815, 2000  
 A: Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venkatesan, K. Nature 407, 87-91, 2000  
 A: Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A: Reference number: AB1000; MUID: 2015755  
 A: Accession: D81055  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-661 <TEXT>  
 A: Cross-references: GB:AE002518; GB:AE002098; NID: g7226928; PIDN: AAF42030.1; PID: g722693  
 A: Experimental source: serogroup B, strain MC58  
 C: Genetics:  
 A: Gene: NM01682  
 C: Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

|                       |        |              |       |            |    |        |      |
|-----------------------|--------|--------------|-------|------------|----|--------|------|
| Query Match           | 31.48; | Score        | 45.5; | DB         | 2; | Length | 661; |
| Best Local Similarity | 34.58; | Pred. No.    | 49;   |            |    |        |      |
| Matches               | 10;    | Conservative | 5;    | Mismatches | 9; | Indels | 5;   |
|                       |        |              |       |            |    | Gaps   | 1;   |

OY    3 LGYHLD----;VSLAFSEISVGAEFNKDD    26  
       :| | : | | : : | | | |  
Db    78 VGLHPPEEGSVVELVFTRLHAGGKFNNKKD    106

Search completed: January 6, 2002, 09:03:01  
Job time: 69 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:04:17 ; Search time 25.81 Seconds  
(without alignments)  
38.355 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID            | Description         |
|------------|-------|-------------|--------|---------------|---------------------|
| 1          | 142   | 97.9        | 559    | 1 CO9_HUMAN   | P02748 homo sapien  |
| 2          | 61.5  | 42.4        | 547    | 1 CO9_HORSE   | P48770 equus caball |
| 3          | 59    | 40.7        | 554    | 1 CO9_RAT     | Q62330 rattus norv  |
| 4          | 57    | 39.3        | 557    | 1 CO9_RABIT   | P48747 oryctolagus  |
| 5          | 49    | 33.8        | 135    | 1 PPOL_ONCMA  | Q08824 oncorhynch   |
| 6          | 49    | 33.8        | 528    | 1 CO9_MOUSE   | P06683 mus musculu  |
| 7          | 46    | 31.7        | 930    | 1 ITTH4_HUMAN | Q14624 h inter-alp  |
| 8          | 44    | 30.3        | 165    | 1 MORE_STNP7  | Q56210 synchococc   |
| 9          | 44    | 30.3        | 560    | 1 KR11_YEAST  | P32893 saccharomyc  |
| 10         | 44    | 30.3        | 668    | 1 F13B_MOUSE  | Q07568 mus musculu  |
| 11         | 44    | 30.3        | 677    | 1 CFPA_TREPA  | Q56336 treponema p  |
| 12         | 44    | 30.3        | 679    | 1 TKTA_CRAPL  | Q42675 craterostig  |
| 13         | 44    | 30.3        | 901    | 1 VGLB_GRCMV  | Q69024 guinea pig   |
| 14         | 44    | 30.3        | 944    | 1 YMH6_YEAST  | Q03631 saccharomyc  |
| 15         | 43.5  | 30.0        | 341    | 1 YG5X_YEAST  | P53336 saccharomyc  |
| 16         | 43    | 29.7        | 305    | 1 YES5_METJA  | Q58850 methanococc  |
| 17         | 43    | 29.7        | 311    | 1 PMG2_YEAST  | Q12008 schizosacch  |
| 18         | 43    | 29.7        | 330    | 1 YDC3_SCHPO  | Q10424 schizosacch  |
| 19         | 43    | 29.7        | 417    | 1 PKG_METJA   | Q58058 methanococc  |
| 20         | 43    | 29.7        | 1070   | 1 P11B_HUMAN  | P42338 homo sapien  |
| 21         | 42.5  | 29.3        | 600    | 1 XKSL_YEAST  | P42826 saccharomyc  |
| 22         | 42    | 29.0        | 238    | 1 ATP6_SPRPN  | Q59954 streptococc  |
| 23         | 42    | 29.0        | 310    | 1 YBRI_YEAST  | P38338 saccharomyc  |
| 24         | 42    | 29.0        | 329    | 1 YX9_BACHD   | Q9K792 bacillus ha  |
| 25         | 42    | 29.0        | 379    | 1 CYB_ASTPE   | Q34318 asterlina pe |
| 26         | 42    | 29.0        | 484    | 1 PURA_MAIZE  | Q24578 zea mays (m  |
| 27         | 42    | 29.0        | 683    | 1 P10I_SCHPO  | P50528 schizosacch  |
| 28         | 42    | 29.0        | 758    | 1 YJ06_YEAST  | P39529 saccharomyc  |
| 29         | 42    | 29.0        | 1052   | 1 BUIB_MOUSE  | Q92102 mus musculu  |
| 30         | 42    | 29.0        | 1877   | 1 PKCS_MOUSE  | Q04592 mus musculu  |
| 31         | 41.5  | 28.6        | 281    | 1 Y246_MYCPN  | P75429 mycoplasma   |
| 32         | 41.5  | 28.6        | 435    | 1 NHG2_PSEPU  | Q53352 pseudomonas  |
| 33         | 41.5  | 28.6        | 790    | 1 YDDB_ECOLI  | P31827 escherichia  |

|    |      |      |      |               |                    |
|----|------|------|------|---------------|--------------------|
| 34 | 41.5 | 28.6 | 986  | 1 EP1B_STRAEP | P30195 staphylococ |
| 35 | 41   | 28.3 | 230  | 1 CRP_RAT     | P48199 rattus norv |
| 36 | 41   | 28.3 | 440  | 1 G3PT_MOUSE  | Q64467 mus musculu |
| 37 | 41   | 28.3 | 481  | 1 6PGD_DROME  | P41572 drosophila  |
| 38 | 41   | 28.3 | 481  | 1 6PGD_DROSI  | P41573 drosophila  |
| 39 | 41   | 28.3 | 560  | 1 EAA5_HUMAN  | Q00341 homo sapien |
| 40 | 41   | 28.3 | 667  | 1 TK12_ECOLI  | P33570 escherichia |
| 41 | 41   | 28.3 | 901  | 1 CR14_MAIZE  | Q24585 zea mays (m |
| 42 | 41   | 28.3 | 921  | 1 ITTH4_PIG   | P79263 sus scrofa  |
| 43 | 41   | 28.3 | 1070 | 1 P11B_RAT    | Q92110 rattus norv |
| 44 | 41   | 28.3 | 2647 | 1 ABP2_HUMAN  | P21333 homo sapien |
| 45 | 40.5 | 27.9 | 250  | 1 Y0U7_CABEL  | P34630 caenorhabdi |

## ALIGNMENTS

| RESULT         | ID  | CO9_HUMAN | STANDARD | PRT | 559 AA. |
|----------------|---|-----------|----------|-----|---------|
| AC             | P02748  |           |          |     |         |
| DT             | 21-JUL-1986 (Rel. 01, Created)  |           |          |     |         |
| DT             | 01-FEB-1996 (Rel. 33, Last sequence update)                             |           |          |     |         |
| DT             | 20-AUG-2001 (Rel. 40, Last annotation update)                           |           |          |     |         |
| DE             | COMPLEMENT COMPONENT C9 PRECURSOR.                                      |           |          |     |         |
| GN             | C9.   |           |          |     |         |
| OS             | Homo sapiens (Human).   |           |          |     |         |
| OC             | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;       |           |          |     |         |
| OC             | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.               |           |          |     |         |
| OX             | NCBI_TaxID=9606;  |           |          |     |         |
| RN             | [1]   |           |          |     |         |
| RP             | SEQUENCE FROM N.A.  |           |          |     |         |
| RX             | MEDLINE=85257464; PubMed=4018030;                                       |           |          |     |         |
| RA             | Stanley K.K., Kocher H.-P., Luzio J.P., Jackson P., Tschoop J.;         |           |          |     |         |
| RT             | "The sequence and topology of human complement component C9.";          |           |          |     |         |
| RL             | EMBO J. 4:375-382(1985).  |           |          |     |         |
| RN             | [2]   |           |          |     |         |
| RP             | SEQUENCE OF 2-559 FROM N.A.   |           |          |     |         |
| RX             | MEDLINE=85063778; PubMed=6095282;                                       |           |          |     |         |
| RA             | Discipio R.G., Gehring M.R., Podack E.R., Kan C.C., Hugli T.E.,         |           |          |     |         |
| RT             | "Nucleotide sequence of cDNA and derived amino acid sequence of human   |           |          |     |         |
| RL             | Proc. Natl. Acad. Sci. U.S.A. 81:7298-7302(1984).                       |           |          |     |         |
| RN             | [3]   |           |          |     |         |
| RP             | SEQUENCE OF 62-159 FROM N.A.  |           |          |     |         |
| RX             | MEDLINE=89118250; PubMed=3219351;                                       |           |          |     |         |
| RA             | Maraziti D., Eggertsen G., Fey G.H., Stanley K.K.;                      |           |          |     |         |
| RT             | "Relationships between the gene and protein structure in human          |           |          |     |         |
| RL             | Biochemistry 27:6529-6534(1988).  |           |          |     |         |
| RN             | [4]   |           |          |     |         |
| RP             | 3D-STRUCTURE MODELLING OF MEMBRANE-SPANNING DOMAIN (MSB).               |           |          |     |         |
| RX             | MEDLINE=90370039; PubMed=2395434;                                       |           |          |     |         |
| RA             | Feitich M.C., Amiguet P., Guy R., Brunner J., Maizel J.V. Jr.,          |           |          |     |         |
| RT             | Tschopp J.;   |           |          |     |         |
| RL             | "Localization and molecular modelling of the membrane-inserted domain   |           |          |     |         |
| RT             | of the ninth component of human complement and perforin.";              |           |          |     |         |
| RL             | Mol. Immunol. 27:589-602(1990).   |           |          |     |         |
| RN             | [5]   |           |          |     |         |
| RP             | CARBOHYDRATE-LINKAGE SITES.   |           |          |     |         |
| RX             | MEDLINE=20020247; PubMed=10551839;                                      |           |          |     |         |
| RA             | Hofsteenge J., Blommestein M., Hess D., Furmanek A., Miroshnichenko O.; |           |          |     |         |
| RT             | "The four terminal components of the complement system are              |           |          |     |         |
| RL             | C-mannosylated on multiple tryptophan residues.";                       |           |          |     |         |
| J. Biol. Chem. | 274:32786-32794(1999).  |           |          |     |         |
| CC             | -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO     |           |          |     |         |
| CC             | BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS          |           |          |     |         |
| CC             | ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.           |           |          |     |         |
| CC             | -1- PTM: THROMBIN CLEAVES FACTOR C9 TO PRODUCE C9A AND C9B.             |           |          |     |         |
| CC             | -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.      |           |          |     |         |
| CC             | -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.                           |           |          |     |         |
| CC             | -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.                             |           |          |     |         |

```

CC CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02176; CAA26117.1; -
DR EMBL; K02766; AAA51889.1; -
DR EMBL; J02833; AAA51890.1; -
DR PIR; A03208; C9HU.
DR HSSP; P01130; IACJ.
DR MIM; 120940; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002172; LDL_receptl_A.
DR InterPro; IPR001862; MAC_perforin.
DR Pfam; PF00057; ldl_receptl_a; 1.
DR Pfam; PF01823; MACPF; 1.
DR PRINTS; PR00764; COMPLEMENTC9.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00457; MACPF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01209; LDLa_1; 1.
DR PROSITE; PS00068; LDLa_2; 1.
DR PROSITE; PS00779; MAC_PERFORIN; 1.
DR PROSITE; PS00092; TSPL; 1.
DR Complement pathway; Complement alternate pathway; Glycoprotein;
DR Plasma; Membrane attack complex; Cytolysis; Transmembrane; Signal;
KW EGF-like domain.
FT SIGNAL 1 21
FT CHAIN 22 559
FT CHAIN 22 265
FT CHAIN 266 559
FT DOMAIN 42 98
FT DOMAIN 99 136
FT TRANSMEM 314 330
FT TRANSMEM 335 354
FT DOMAIN 506 540
FT SITE 265 266
FT DISULFID 107 112
FT DISULFID 107 125
FT DISULFID 119 134
FT DISULFID 254 255
FT DISULFID 380 405
FT DISULFID 510 526
FT DISULFID 513 528
FT CARBOHYD 48 539
FT CARBOHYD 51 51
FT CARBOHYD 277 277
FT CARBOHYD 415 415
FT CONFLICT 43 43
FT CONFLICT 314 314
FT CONFLICT 417 417
SQ SEQUENCE 559 AA; 63173 MW; 7403F6AD77B3ECE1 CRC64;

Query Match 97.9%; Score 142; DB 1; Length 559;
Best Local Similarity 96.3%; Pred. No. 1.8e-14;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KCLGYHLDVSLAFSEISVGAEEFKKDDC 27
Db < 379 KCLGYHLDVSLAFSEISVGAEEFKKDDC 405

```

```

RESULT 2
ID CO9_HORSE STANDARD; PRT; 547 AA.
AC P48770;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE COMPLEMENT COMPONENT C9 PRECURSOR.
GN C9.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_Taxid-9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95325619; PubMed-7541424;
RA Tomlinson S., Wang Y., Ueda E., Esser A.F.;
RT "Chimeric horse/human recombinant C9 proteins identify the amino acid
RT sequence in horse C9 responsible for restriction of hemolysis.";
RL J. Immunol. 155:436-444(1995).
CC CC -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO
CC BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS
CC ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.
CC -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19381; AAB16820.1; -
DR HSSP; P01130; IACJ.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002172; LDL_receptl_A.
DR InterPro; IPR001862; MAC_perforin.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00057; ldl_receptl_a; 1.
DR Pfam; PF01823; MACPF; 1.
DR PRINTS; PR00764; COMPLEMENTC9.
DR SMART; SM00001; EGF_1like; 1.
DR SMART; SM00019; EGF_1like; 1.
DR SMART; SM00457; MACPF; 1.
DR SMART; SM00457; MACPF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE.NEG.
DR PROSITE; PS01209; LDLa_1; 1.
DR PROSITE; PS00068; LDLa_2; 1.
DR PROSITE; PS00779; MAC_PERFORIN; 1.
DR PROSITE; PS00092; TSPL; 1.
DR Complement pathway; Complement alternate pathway; Glycoprotein;
DR Plasma; Membrane attack complex; Cytolysis; Transmembrane; Signal;
KW EGF-like domain.
FT SIGNAL 1 21
FT CHAIN 22 547
FT CHAIN 22 98
FT CHAIN 99 136
FT DOMAIN 314 330
FT TRANSMEM 335 354
FT TRANSMEM 506 540
FT DOMAIN 101 112
FT DISULFID 107 112
FT DISULFID 107 125
FT DISULFID 119 134
FT DISULFID 380 404
FT DISULFID 510 526
FT DISULFID 513 528
FT DISULFID 513 528

```

```

FT DISULFID 530 539 BY SIMILARITY.
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 547 AA; 62014 MW; 75E5FE961DE873B6 CRC64;

Query Match 42.4%; Score 61.5; DB 1; Length 547;
Best Local Similarity 44.4%; Pred. No. 0.062;
Matches 12; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

OY 1 KCICGHLDVSLAFSEISVGAEFNKDDC 27
Db 379 RCLGFNLDLSLK-DKYEVTAKIDKNDK 404

RESULT 3
CO9_RAT 3 STANDARD; PRT: 554 AA.
AC 062930;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE COMPLEMENT COMPONENT C9 PRECURSOR.
GN C9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RC MEDLINE=97355567; PubMed=9212048;
RA Lassiter H.A., Walz B.M., Wilson J.L., Jung E., Calisi C.R.,
RA Goldsmith L.J., Wilson R.A., Morgan B.P., Feldhoff R.C.,
RT "The administration of complement component C9 enhances the survival
RT of neonatal rats with Escherichia coli sepsis."
RL Pediatr. Res. 42:128-136(1997).
CC -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO
CC BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS
CC ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.
CC -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U49071; AAB38023.1; -
CC HSSP: P01130; ILDL.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR002172; LDL-recept.A.
CC InterPro: IPR001862; MAC_perforin.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00057; ldl_receptl_a; 1.
CC Pfam: PF01823; MACPF; 1.
CC PRINTS: PR00764; COMPLEMENTC9.
CC SMART: SM00192; LDLa; 1.
CC SMART: SM00457; MACPF; 1.
CC SMART: SM00209; TSP1; 1.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; FALSE_NEG.
CC PROSITE: PS01209; ILDLA_1; 1.
CC PROSITE: PS50068; ILDLA_2; 1.
CC PROSITE: PS00279; MAC_PERFORIN; 1.
CC PROSITE: PS50092; TSP1; 1.

```

```

KW Complement pathway; Complement alternate pathway; Glycoprotein;
KW Plasma; Membrane attack complex; Cytolysis; Transmembrane; Signal;
KW EGF-like domain.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 554 COMPLEMENT COMPONENT C9.
FT DOMAIN 40 96 TSP TYPE-1.
FT DOMAIN 97 134 LDL-RECEPTOR CLASS A.
FT TRANSMEM 316 332 POTENTIAL.
FT TRANSMEM 337 356 POTENTIAL.
FT DOMAIN 515 549 EGF-LIKE.
FT SITE 267 268 CLEAVAGE (BY THROMBIN).
FT DISULFID 99 110 BY SIMILARITY.
FT DISULFID 105 123 BY SIMILARITY.
FT DISULFID 117 132 BY SIMILARITY.
FT DISULFID 382 413 BY SIMILARITY.
FT DISULFID 519 535 BY SIMILARITY.
FT DISULFID 522 537 BY SIMILARITY.
FT DISULFID 539 548 BY SIMILARITY.
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 554 AA; 62280 MW; 9C885F76A1275649 CRC64;

Query Match 40.7%; Score 59; DB 1; Length 554;
Best Local Similarity 42.4%; Pred. No. 0.16;
Matches 14; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 1 KCICGHLDVSL-----AFSEISVGAEFNKDDC 27
Db 381 RCLGFNLDVSLYTPLOTALGPSLTJANVHNSDC 413

RESULT 4
CO9_RABIT 4 STANDARD; PRT: 557 AA.
AC P48747;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE COMPLEMENT COMPONENT C9 PRECURSOR.
GN C9.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-34.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;
RC MEDLINE=95181293; PubMed=753152;
RA Huesler T., Lockert D.H., Kaufman K.M., Sodez J.M., Sims P.J.;
RT "Chimers of human complement C9 reveal the site recognized by
RT complement regulatory protein CD59."
J. Biol. Chem. 270:3483-3486(1995).
CC -1- FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO
CC BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS
CC ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRANE CHANNELS.
CC -1- PTM: THROMBIN CLEAVES FACTOR C9 TO PRODUCE C9A AND C9B.
CC -1- SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U20055; AAC48459.1; -
CC HSSP: P01130; IAUJ.

```

| RESULT | 5  |           |      |         |
|--------|--|-----------|------|---------|
| ID     | PPOL_ONCMA   | STANDARD: | PRT: | 135 AA. |
| AC     | 000824;  |           |      |         |
| DT     | 01-OCT-1994 (Rel. 30, Created)   |           |      |         |
| DT     | 01-FEB-1995 (Rel. 31, Last sequence update)                                |           |      |         |
| DT     | 20-AUG-2001 (Rel. 40, Last annotation update)                              |           |      |         |
| DE     | POLY (ADP-RIBOSE) POLYMERASE (EC 2.4.2.30) (PAP)                           |           |      |         |
| DE     | RIBOSYLTRANSFERASE (POLY(ADP-RIBOSE) SYNTHETASE) (FRAGMENT).               |           |      |         |
| GN     | ADPRT.   |           |      |         |
| OS     | Oncorhynchus mason (Cherry salmon) (Masu salmon).                          |           |      |         |
| OC     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;          |           |      |         |
| OC     | Actinopterygii; Neopterygii; Teleostei; Euteleostei;                       |           |      |         |
| OC     | Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.              |           |      |         |
| OX     | NCBI_TaxID=8020;   |           |      |         |
| RN     | [1]  |           |      |         |
| RP     | SEQUENCE FROM N.A.   |           |      |         |
| RX     | MEDLINE=93277538; PubMed=8503897;  |           |      |         |
| RA     | Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,              |           |      |         |
| RA     | Miwa M.,   |           |      |         |
| RT     | "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)      |           |      |         |
| RT     | polymerase from <i>Xenopus laevis</i> and cherry salmon using heterologous |           |      |         |
| RT     | oligonucleotide consensus sequences."                                      |           |      |         |
| RL     | Biochem. Biophys. Res. Commun. 193:119-125(1993).                          |           |      |         |

```

CC -1- PROTEIN. POLY(ADP-RIBOSE)] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC -2- FUNCTION BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -3- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-RIBOSY)[n-1]-ACCEPTOR =
CC NICOTINAMIDE + (ADP-D-RIBOSY)[n-1]-ACCEPTOR.
CC -4- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC -5- SUBCELLULAR LOCATION: NUCLEAR.
CC -6- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC AND BRAIN. LOW IN LIVER.
CC -7- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF. AND
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC -8- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13809; BAA02965.1; -.
DR PIR: P04994; P04994.
DR HSSP: P26446; 4PAX.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR001510; Znf-PARP.
DR Pfam: PF00644; PARP. 1.
DR PROSITE: PS00347; PARP_ZN_FINGER_1; PARTIAL.
DR PROSITE: PS50064; PARP_ZN_FINGER_2; PARTIAL.
KW Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW ADP-ribosylation; zinc-finger; zinc.
FT NON_TER 1 1
FT DOMAIN <1 >135 NAD-BINDING.
FT ACT_SITE 135 135 BY SIMILARITY.
FT NON_TER 135 135
SEQUENCE 135 AA; 15411 MW; A608E98890FA27C Cn664.

```

```

Query Match      33.88;  Score 49;  DB 1;  Length 135;
Best Local Similarity 50.0%;
Matches 10;  Conservative 4;  Mismatches 6;  Indels 0;  Gaps 0;
OY      7  LDVSLAFSEISVGAEPKKDD 26
      ||::||: || || ||
Db      11 LDIEAVSLKGAEDNKKD 30

RESULT 6
CO9_MOUSE
ID      CO9_MOUSE      STANDARD;      PRT;      528 AA.
AC      P06683;
DT      01-JAN-1988 (Rel. 06, Created)
DT      01-JAN-1988 (Rel. 06, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      COMPLEMENT COMPONENT C9 (FRAGMENT).
GN      C9.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=88004404; PubMed=2443347;
RA      Stanley R.K., Herz J.;
RT      "Topological mapping of complement component C9 by recombinant DNA
RT      techniques suggests a novel mechanism for its insertion into target
RT      membranes.";
```

```

RL  EMBO J. 6:1951-1957(1987).
CC  - FUNCTION: C9 IS THE FINAL COMPONENT OF THE COMPLEMENT SYSTEM TO
CC  BE ADDED IN THE ASSEMBLY OF THE MEMBRANE ATTACK COMPLEX. IT IS
CC  ABLE TO ENTER LIPID BILAYERS, FORMING TRANSMEMBRAN CHANNELS.
CC  - SIMILARITY: TO COMPLEMENT FACTORS C6, C7, C8, AND TO PERFORIN.
CC  - SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC  - SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC  - SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X05475; CAA29038.1; -
DR  PIR: A29677; A29677.
DR  MGD: MGI:1098282; C9.
DR  InterPro: IPR000561; EGF-like.
DR  InterPro: IPR002172; LDL_recept_L.
DR  InterPro: IPR001862; MAC_perforin.
DR  InterPro: IPR000884; TSP1.
DR  Pfam: PF00057; 1dl_recept_a; 1.
DR  Pfam: PF01823; MACPF; 1.
DR  Pfam: PF00090; tsp_1; 1.
DR  SMART: SM00192; LDla; 1.
DR  SMART: SM00457; MACPF; 1.
DR  SMART: SM00209; TSP1; 1.
DR  PROSITE: PS00022; EGF_1; 1.
DR  PROSITE: PS01186; EGF_2; FALSE_NEG.
DR  PROSITE: PS01209; LDLRA_1; 1.
DR  PROSITE: PS50068; LDLRA_2; 1.
DR  PROSITE: PS00279; MAC_PERFORIN; 1.
DR  PROSITE: PS50092; TSP1; 1.
DR  PROSITE: PS50092; TSP1; 1.
KW  Complement pathway; Complement alternate pathway; Glycoprotein;
KW  Plasma; Membrane attack complex; Cytolysis; Transmembrane;
KW  EGF-like domain.
FT  NON_TER 1 76 TSP TYPE-1.
FT  DOMAIN 1 20 1 LDL-RECEPTOR CLASS A.
FT  TRANSMEM 292 114 POTENTIAL.
FT  TRANSMEM 313 308 POTENTIAL.
FT  DOMAIN 489 523 EGF-LIKE.
FT  DISULFID 79 90 BY SIMILARITY.
FT  DISULFID 85 103 BY SIMILARITY.
FT  DISULFID 97 112 BY SIMILARITY.
FT  DISULFID 358 387 BY SIMILARITY.
FT  DISULFID 493 511 BY SIMILARITY.
FT  DISULFID 496 509 BY SIMILARITY.
FT  DISULFID 513 522 BY SIMILARITY.
FT  CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 528 AA; 60020 MW; DID096057D10E5C5 CRC64;

```

```

Query Match 33.8%; Score 49; DB 1; Length 528;
Best Local Similarity 33.3%; Pred. No. 5.3;
Matches 10; Conservative 6; Mismatch 10; Indels 4; Gaps 1;

```

```

OY 2 CAGYHADVSLA---FSEISVGAEFNKDDC 27
DB 358 CAGFNMDLRIPIQDDLDKASVTSVADGC 387

```

```

RESULT 7
ID ITI4_HUMAN STANDARD: PRT: 930 AA.
AC 014624: 015135: 09U054;
DT 15-JUL-1998 (Rel. 36; Created)
DT 20-AUG-2001 (Rel. 40; Last sequence update)

```

```

DT 20-AUG-2001 (Rel. 40; Last annotation update)
DE INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN H4 PRECURSOR (ITI HEAVY
DE CHAIN H4) (INTER-ALPHA-TRYPSIN INHIBITOR FAMILY HEAVY CHAIN-RELATED
DE PROTEIN) (ITIRP) (PLASMA KALLIKREIN SENSITIVE GLYCOPROTEIN 120) (RK-
DE 120) (GPI20) [CONTAINS: GP57].
GN ITI4 OR IHRP OR ITIH1 OR PK120.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95293915; PubMed=7775381;
RA Saguichi K., Tobe T., Hashimoto K., Sano Y., Nakano Y., Miura N.-H.,
RA Tomita M.;
RT Cloning and characterization of cDNA for inter-alpha-trypsin
RT inhibitor family heavy chain-related protein (IHRP), a novel human
RT plasma glycoprotein."
RL J. Biochem. 117:14-18(1995).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95104473; PubMed=7805892;
RA Nishimura H., Kakizaki I., Muta T., Sasaki N., Pu P.X., Yamashita T.,
RA Nagasawa S.;
RT cDNA and deduced amino acid sequence of human PK-120, a plasma
RT kallikrein-sensitive glycoprotein."
RL FEBS Lett. 357:207-211(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96389995; PubMed=8797089;
RA Saguichi K., Tobe T., Hashimoto K., Nagasaki Y., Oda E., Nakano Y.,
RA Miura N.H., Tomita M.;
RT Isolation and characterization of the human inter-alpha-trypsin
RT inhibitor family heavy chain-related protein (IHRP) gene (ITIH1).";
RL J. Biochem. 119:898-905(1996).
RN [4]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Plasma;
RX MEDLINE=95332266; PubMed=7541790;
RA Choi-Miura N.-H., Sano Y., Oda E., Nakano Y., Tobe T., Yanagishita T.,
RA Tanigawa M., Katagiri T., Tomita M.;
RT Purification and characterization of a novel glycoprotein which has
RT significant homology to heavy chains of inter-alpha-trypsin inhibitor
RT family from human plasma."
RL J. Biochem. 117:400-407(1995).
RN [5]
RP FUNCTION: MAY BE INVOLVED IN ACUTE PHASE REACTIONS.
CC - FUNCTION: MAY BE INVOLVED IN ACUTE PHASE REACTIONS.
CC - TISSUE SPECIFICITY: LIVER-SPECIFIC.
CC - INDUCTION: LEVELS INCREASE SIGNIFICANTLY IN THE SERA OF PATIENTS
CC AFTER DIFFERENT SURGICAL TRAUMA.
CC - PTM: CLEAVED BY PLASMA KALLIKREIN TO YIELD 100- AND 35-KDA
CC FRAGMENTS, AND THE RESULTING 100-KDA FRAGMENT IS FURTHER CONVERTED
CC TO A 70-KDA FRAGMENT.
CC - PTM: APPEARS TO BE BOTH N- AND O-GLYCOSYLATED.
CC - SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC - SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D38595; BAA07602.1; -
DR EMBL: D38535; BAA07536.1; -
DR EMBL: U43163; AAD05198.1; -
DR EMBL: U42015; AAD05198.1; JOINED.
DR EMBL: U42016; AAD05198.1; JOINED.
DR EMBL: U43155; AAD05198.1; JOINED.
DR EMBL: U43156; AAD05198.1; JOINED.

```

```

DR EMBL: U43157; AAD05198.1; JOINED.
DR EMBL: U43158; AAD05198.1; JOINED.
DR EMBL: U43159; AAD05198.1; JOINED.
DR EMBL: U43160; AAD05198.1; JOINED.
DR EMBL: U43161; AAD05198.1; JOINED.
DR EMBL: U43162; AAD05198.1; JOINED.
DR MIM: 600564; -.
DR InterPro: IPR002035; VFMA.
DR Pfam: PF00092; VMA; 1.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS50234; VFMA; 1.
KW Serine protease inhibitor; Repeat; Signal; Multigene family;
KW Glycoprotein.
FT SIGNAL 1 28
FT CHAIN 29 661
FT PROPEP 662 668
FT CHAIN 689 930
FT DOMAIN 272 432
FT CARBOHYD 81 81
FT CARBOHYD 207 207
FT CARBOHYD 517 517
FT CARBOHYD 577 577
FT CARBOHYD 666 666
FT CARBOHYD 701 701
FT CARBOHYD 702 702
FT DISULFID 747 925
FT CONFLICT 85 85
FT CONFLICT 85 85
FT CONFLICT 114 114
FT CONFLICT 207 207
FT CONFLICT 221 221
FT CONFLICT 307 307
FT CONFLICT 322 322
FT CONFLICT 816 817
SQ SEQUENCE 930 AA; 103358 MW; 0E1929065FE4EB6A0 CRC64;

```

```

Query Match
Best Local Similarity 31.7%; Score 46; DB 1; Length 930;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 2 CIGYHLDVSLAFSEI5VGAEPNKDD 26
Db 414 CIGYHLDVSLAFSEI5VGAEPNKDD 427

RESULT 9
ID MOAE_SYNP7 STANDARD; PRT; 165 AA.
AC 056210;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE MOLYBDOTERIN [MPT] CONVERTING FACTOR, SUBUNIT 2 (MOLYBDENUM COFACTOR
DE BIOSYNTHESIS PROTEIN E) (MOLYBDOTERIN CONVERTING FACTOR LARGE
DE SUBUNIT).
GN MOAE.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubio L.M., Flores E., Herrero A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: CONVERSION OF MOLYBDOTERIN PRECURSOR 2 INTO
CC MOLYBDOTERIN REQUIRES TRANSFER OF TWO SULFUR ATOMS TO PRECURSOR 2
CC (TO GENERATE THE DITHIOLENE GROUP). THIS IS CATALYZED BY THE
CC CONVERTING FACTOR COMPOSED OF A SMALL AND LARGE SUBUNIT (BY
CC SIMILARITY).
CC -I- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS.
CC -I- SUBUNIT: HETERODIMER OF MOAD AND MOAE.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99625; CAA67947.1; -.
DR InterPro: IPR003448; Moea.
DR Pfam: PF02391; Moea; 1.
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 165 AA; 18712 MW; AC016984E1D587FB CRC64;

```

```

Query Match
Best Local Similarity 30.3%; Score 44; DB 1; Length 165;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 2 CIGYHLDVSLAFSEI5VGAEPNKDD 26
Db 7 CDRHIELSLAPRPLSAAEFCDD 31

```

```

RESULT 9
ID KRI1_YEAST STANDARD; PRT; 560 AA.
AC P32893;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BETA-GLUCAN SYNTHESIS-ASSOCIATED PROTEIN KREI1 (KILLER TOXIN-
DE RESISTANCE PROTEIN 11).
GN KREI1 OR YGR166M.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93216084; PubMed=8462845;
RA Brown J.L., Kosaczka Z., Jiang B., Bussey H.;
RT "A mutational analysis of killer toxin resistance in Saccharomyces
RT cerevisiae identifies new genes involved in cell wall (1-->6)-beta-
RT glucan synthesis."
RL Genetics 133:837-849(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII."
RL Yeast 13:1077-1090(1997).
CC -I- FUNCTION: INVOLVED IN CELL WALL (1-->6)-BETA-GLUCAN SYNTHESIS.
CC KREI1 PROBABLY INTERACTS WITH KRB6. THEIR ROLE PRECEDES THAT OF
CC -I- PATHWAY: BIOSYNTHESIS OF BETA-GLUCANS.
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L10667; AA34727.1; -.
DR EMBL: Z72951; CAA97191.1; -.
DR PIR: S30784; S30784.
DR PIR: S42158; S42158.

```





```

DR EMBL:Z46647; CAAB86608.1; -.
DR HSP: P23254; IAT0.
DR InterPro: IPR000360; Transketolase.
DR Pfam: PF00456; transketolase_1.
DR PROSITE: PS00801; TRANSKETOLASE_1; 1.
DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
KW transferase; Thiamine pyrophosphate; Multigene family.
SQ SEQUENCE 679 AA; 73130 MW; 1109092E136A345B CRC64;

Query Match
Best Local Similarity 30.3%; Score 44; DB 1; Length 679;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0

Oy 9 VSIAFSEISGAFENKDD 26
Db 139 VGLALAEKHIGAKYKPD 156
||| : || : ||| |
-----

RESULT 13
VGLB-GPCMV STANDARD; PRT; 901 AA.
ID AC 069024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GLYCOPROTEIN B PRECURSOR.
GN GB OR UL55.
OS guinea pig cytomegalovirus (strain 2212 / ATCC VR682) (GPCMV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muricamegaloivirus.
NCBI_TaxID=103920;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94279138; PubMed=8009831;
RA Schieles M.R.;
RT "Cloning and characterization of the guinea pig cytomegalovirus
RL glycoprotein B gene."
VL Virology 202:173-185(1994).
--1 SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: L25706; AAA43831.1; -.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B_1.
DR Prodom: PD000693; Glycoprot_B_1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 901 POTENTIAL.
FT DOMAIN ? 713 GLYCOPROTEIN B.
FT TRANSMEM 714 734 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 737 757 POTENTIAL.
FT DOMAIN 758 901 POTENTIAL.
FT CARBOHYD 53 53 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAG. . .) (POTENTIAL).

```

FT CARBOHYD 623 623 N-LINKED (GLCNAC.. ) (POTENTIAL).  
SQ SEQUENCE 901 AA; 102239 MW; 43ABCA/D45A197AA CRC64;

Query Match 30.3%; Score 44; DB 1; Length 901;  
Best Local Similarity 34.6%; Pred. No. 55;  
Matches 9; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

OY 1 KCLGYHLVSLAFSEISVGAFFNKDD 26  
DB 815 KDLSDVADTRVSSSSAGAGDFNEED 840

## RESULT 14

YMH6\_YEAST  
ID YMH6\_YEAST STANDARD; PRT; 944 AA.  
AC Q03631;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PUTATIVE 107.6 KDA TRANSCRIPTIONAL REGULATORY PROTEIN IN CPR3-HMG1  
DE INTERGENIC REGION.  
GN YML076C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Brown D., Bowman S., Barrell B.G., Rajandream M.A.;  
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BIOMOLECULAR  
CC CLUSTER DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z46373; CAA86502.1; -  
DR SGD: S0004541; YML076C.  
DR InterPro: IPR001138; ZN2\_Cy6\_fungal.  
DR SMART: SM00066; GAL4; 1.  
DR PROSITE: PS00463; ZN2\_Cy6\_FUNGAL\_1; 1.  
DR PROSITE: PS50048; ZN2\_Cy6\_FUNGAL\_2; FALSE\_NEG.  
KW Hypothetical protein; Transcription regulation; DNA-binding; Zinc;  
KW Nuclear protein; Metal-binding.  
FT DNA\_BIND 76 109 ZN(2)-CYS(6), FUNGAL-TYPE.  
SQ SEQUENCE 944 AA; 107560 MW; 020A56745DF52CCC CRC64;

Query Match 30.3%; Score 44; DB 1; Length 944;  
Best Local Similarity 35.0%; Pred. No. 58;  
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 7 LDVSLAFSEISVGAFFNKDD 26  
DB 1 MDPIAITGVAVGKEINNND 20

## RESULT 15

YG5X\_YEAST  
ID YG5X\_YEAST STANDARD; PRT; 341 AA.  
AC P53336;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 38.5 KDA PROTEIN IN BGL2-ZUO1 INTERGENIC REGION.

GN YGR283C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1679;  
RX MEDLINE=97245295; PubMed=9090054;  
RA Volckaert G., Voet M., Robben J.;  
RT "Sequence analysis of a near-subtelomeric 35.4 kb DNA segment on the  
RT right arm of chromosome VII from Saccharomyces cerevisiae carrying  
RT the MAL1 locus reveals 15 complete open reading frames, including  
RT ZUO1, BGL2 and BIO2 genes and an ABC transporter gene.";  
RL Yeast 13:251-259(1997).  
CC -1- SIMILARITY: TO YEAST YMR310C.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: Z73068; CAA97314.1; -  
DR SGD: S0003515; YGR283C.  
DR InterPro: IPR003750; DUF171.  
DR Pfam: PF02598; DUF171; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 341 AA; 38546 MW; 3C91F8EF9C24EFC9 CRC64;

Query Match 30.0%; Score 43.5; DB 1; Length 341;  
Best Local Similarity 28.6%; Pred. No. 25;  
Matches 10; Conservative 8; Mismatches 8; Indels 9; Gaps 1;

OY 1 KCLGYHLVSLAFSEISVG-----AEPNKDD 26  
DB 216 KPLGYHMASTLNEVSEGYTKIYVWNSGDRHYDE 250

Search completed: January 6, 2002, 09:11:32  
Job time: 435 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:03:07 ; Search time 62.25 Seconds  
(without alignments)  
63.443 Million cell updates/sec

Title: US-09-020-393b-18  
Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 59    | 40.7        | 567    | 11    | Q62957 rattus norv |
| 2          | 52    | 35.9        | 303    | 2     | Q9E231 neisseria 9 |
| 3          | 52    | 35.9        | 326    | 2     | 051948 neisseria 9 |
| 4          | 52    | 35.9        | 326    | 2     | 0924K1 neisseria 9 |
| 5          | 52    | 35.9        | 326    | 2     | 051944 neisseria 9 |
| 6          | 52    | 35.9        | 327    | 2     | 0924L5 neisseria 9 |
| 7          | 52    | 35.9        | 327    | 3     | 013834 schizosacch |
| 8          | 47    | 32.4        | 241    | 5     | 0917L2 drosophila  |
| 9          | 47    | 32.4        | 278    | 10    | Q9SSX1 oryza sativ |
| 10         | 47    | 32.4        | 367    | 10    | Q9SNG2 oryza sativ |
| 11         | 47    | 32.4        | 391    | 5     | 09V3Z2 drosophila  |
| 12         | 46.5  | 32.1        | 1002   | 5     | 045247 caenorhabd  |
| 13         | 46    | 31.7        | 203    | 12    | Q9YMI5 lymantria d |
| 14         | 46    | 31.7        | 254    | 10    | Q9XEF0 arabidopsis |
| 15         | 46    | 31.7        | 292    | 2     | 067397 aquifex aeo |
| 16         | 46    | 31.7        | 357    | 12    | 067692 garlic viru |
| 17         | 46    | 31.7        | 376    | 2     | 087140 vibrio chol |
| 18         | 46    | 31.7        | 377    | 2     | 034236 vibrio chol |
| 19         | 46    | 31.7        | 644    | 4     | Q9P190 homo sapien |

|    |      |      |      |    |        |                    |
|----|------|------|------|----|--------|--------------------|
| 20 | 46   | 31.7 | 1372 | 5  | P91526 | P91526 caenorhabd  |
| 21 | 46   | 31.7 | 2144 | 5  | 062218 | 062218 caenorhabd  |
| 22 | 45.5 | 31.4 | 661  | 2  | Q9JY97 | Q9JY97 neisseria m |
| 23 | 45.5 | 31.4 | 661  | 2  | Q9JY79 | Q9JY79 neisseria m |
| 24 | 45   | 31.0 | 358  | 2  | 087162 | 087162 vibrio chol |
| 25 | 45   | 31.0 | 440  | 1  | Q9HOM6 | Q9HOM6 halobacteri |
| 26 | 45   | 31.0 | 732  | 5  | 018367 | 018367 caenorhabd  |
| 27 | 45   | 31.0 | 895  | 10 | Q9LX29 | Q9LX29 arabidopsis |
| 28 | 45   | 31.0 | 1189 | 5  | 018558 | 018558 aedes aegy  |
| 29 | 45   | 31.0 | 2524 | 5  | 09GPAS | 09GPAS branchiost  |
| 30 | 44.5 | 30.7 | 623  | 10 | 081082 | 081082 allium cepa |
| 31 | 44.5 | 30.7 | 909  | 5  | 017429 | 017429 caenorhabd  |
| 32 | 44   | 30.3 | 71   | 6  | 029226 | 029226 sus scrofa  |
| 33 | 44   | 30.3 | 160  | 6  | Q9GLE7 | Q9GLE7 sus scrofa  |
| 34 | 44   | 30.3 | 241  | 5  | 062460 | 062460 caenorhabd  |
| 35 | 44   | 30.3 | 261  | 2  | 09CGJ1 | 09CGJ1 lactococcus |
| 36 | 44   | 30.3 | 311  | 10 | Q9LJW1 | Q9LJW1 arabidopsis |
| 37 | 44   | 30.3 | 380  | 2  | Q9FCU8 | Q9FCU8 lactobacill |
| 38 | 44   | 30.3 | 423  | 8  | Q9B896 | Q9B896 schistosoma |
| 39 | 44   | 30.3 | 424  | 8  | Q9B8Z8 | Q9B8Z8 schistosoma |
| 40 | 44   | 30.3 | 535  | 2  | 09JZ11 | 09JZ11 neisseria m |
| 41 | 44   | 30.3 | 535  | 2  | 09JU07 | 09JU07 neisseria m |
| 42 | 44   | 30.3 | 571  | 2  | Q9LAP0 | Q9LAP0 staphylococ |
| 43 | 44   | 30.3 | 659  | 2  | Q9JTS0 | Q9JTS0 neisseria m |
| 44 | 44   | 30.3 | 659  | 2  | Q9JTR1 | Q9JTR1 neisseria m |
| 45 | 44   | 30.3 | 662  | 2  | Q99UD4 | Q99UD4 staphylococ |

#### ALIGNMENTS

RESULT 1  
ID 062957 PRELIMINARY; PRT; 567 AA.  
AC 062957;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE COMPLEMENT COMPONENT C9 PRECURSOR.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
ON NCBI\_TaxID=10116;  
RX (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE DAWLEY;  
RA Hinchliffe S.J., den Berg C.W., Rushmere N.K., Morgan B.P.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR CLASS A (LDLRA) DOMAIN.  
DR EMBL: U52948; AAA96528.1; -.  
DR HSSP: P01130; LDL.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002172; LDL\_recept\_A.  
DR InterPro: IPR001862; MAC\_perforin.  
DR InterPro: IPR000884; TSP1.  
DR Pfam: PF00057; ldl\_recept\_a; 1.  
DR Pfam: PF01823; MACPF; 1.  
DR Pfam: PF00090; TSP\_1; 1.  
DR PRINTS: PR00764; COMPLEMENTC9.  
DR SMART: SM00192; LDLa; 1.  
DR SMART: SM00457; MACPF; 1.  
DR SMART: SM00209; TSP1; 1.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE: PS01209; LDLRA\_1; 1.  
DR PROSITE: PS0068; LDLRA\_2; 1.  
DR PROSITE: PS00279; MAC\_PERFORIN; 1.  
DR PROSITE: PS00092; TSP1; 1.  
DR GlycoProtex: Signal.  
FT SIGNAL 1 33 POTENTIAL.  
FT CHAIN 34 567 COMPLEMENT COMPONENT C9.  
SQ SEQUENCE 567 AA; 63762 MW; DBA9C728490AE68A CRC64;

```

Query Match
Best Local Similarity 40.7%; Score 59; DB 11; Length 567;
Matches 14; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 1 KCLGYHLDVSLAFSEISVGAEPN 27
    |||::|||::| 1 1 1 1 1
Db 394 KCLGFNLDVSLYTPLOTALTEGSLTAVNHSDC 426

RESULT 2
OQ9EZ31 PRELIMINARY; PRT; 303 AA.
AC OQ9EZ31;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PORIN (FRAGMENT).
GN POR.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20553344; PubMed=11101576;
RA Viscidi R.P., Demma J.C., Gu J., Zenilman J.;
RT "Comparison of Sequencing of the por Gene and Typing of the opa Gene
RT for Discrimination of Neisseria gonorrhoeae Strains from Sexual
RT Contacts."
RL J. Clin. Microbiol. 38:4430-4438(2000).
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
DR EMBL: AF304405; AAC24903.1; -
DR InterPro: IPR001702; Gram_neg_porin.
DR Pfam: PF00267; Gram-ve_porins. 1.
DR PRINTS: PR00182; ECOLNEIPORIN.
DR PROSITE: PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Porin; Signal; Transmembrane.
FT NON_TER 1
FT NON_TER 303
FT SEQUENCE 303 AA; 33187 MW; 08D22F4EF697F7EA CRC64;

Query Match
Best Local Similarity 35.9%; Score 52; DB 2; Length 303;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEPN 23
    |||::|||::| 1 1 1 1 1
Db 255 KGLVYHADLSNDYDVVGAEYD 277

RESULT 3
OQ51948 PRELIMINARY; PRT; 326 AA.
AC OQ51948;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER MEMBRANE PROTEIN I PRECURSOR.
GN POR.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98258030; PubMed=9595666;
RA Cooke S.J., Jolley K., Ison C.A., Young H., Heckels J.E.;
RT "Naturally occurring isolates of Neisseria gonorrhoeae, which display
RT anomalous serovar properties, express PIA/PIB hybrid porins, deletions

```

```

RT in PIB or novel PIA molecules."
RL EEMS Microbiol. Lett. 162:75-82(1998).
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
DR EMBL: AF015122; AAC36568.1; -
DR InterPro: IPR001702; Gram_neg_porin.
DR Pfam: PF00267; Gram-ve_porins. 1.
DR PRINTS: PR00182; ECOLNEIPORIN.
DR PROSITE: PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Porin; Signal; Transmembrane.
FT SIGNAL 1
FT CHAIN 20 326 POTENTIAL.
FT SEQUENCE 326 AA; 35558 MW; 3D566D777D7CA015 CRC64;

Query Match
Best Local Similarity 35.9%; Score 52; DB 2; Length 326;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEPN 23
    |||::|||::| 1 1 1 1 1
Db 268 KGLVYHADLSNDYDVVGAEYD 290

RESULT 4
OQ9ZAK1 PRELIMINARY; PRT; 326 AA.
ID OQ9ZAK1
AC OQ9ZAK1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER MEMBRANE PROTEIN I PRECURSOR.
GN POR.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=4846;
RA Pudys T.C., Maclean I.W., Simonsen J.N., Njagi E.N., Kimani J.,
RA Brunham R.C., Plummer F.A.;
RT "Genetic Diversity and Mosaicism at the por Locus of Neisseria
RT gonorrhoeae."
RT Submitted (SEP-1998) to the EMBL/GenBank/DDJ databases.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
DR EMBL: AF090824; AAC79493.1; -
DR InterPro: IPR001702; Gram_neg_porin.
DR Pfam: PF00267; Gram-ve_porins. 1.
DR PRINTS: PR00182; ECOLNEIPORIN.
DR PROSITE: PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Porin; Signal; Transmembrane.
FT SIGNAL 1
FT CHAIN 20 326 POTENTIAL.
FT SEQUENCE 326 AA; 35611 MW; 5F0CC17B6A8FA323 CRC64;

Query Match
Best Local Similarity 35.9%; Score 52; DB 2; Length 326;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEPN 23
    |||::|||::| 1 1 1 1 1
Db 268 KGLVYHADLSNDYDVVGAEYD 290

RESULT 5
OQ51944 PRELIMINARY; PRT; 326 AA.
ID OQ51944
AC OQ51944;
DT 01-JUN-1998 (TREMBLrel. 06, Created)

```

```

DT 01-JUN-1998 (TReMBLrel. 06, last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)
DE OUTER MEMBRANE PROTEIN I PRECURSOR.
GN POR.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=31493;
RX MEDLINE=98258030; PubMed=9595666;
RA Cooke S.J., Jolley K., Jison C.A., Young H., Heckels J.E.;
RT "Naturally occurring isolates of Neisseria gonorrhoeae, which display
RT anomalous serovar properties, express PIA/PilB hybrid porins, deletions
RT in pilB or novel pilA molecules."
RL FEMS Microbiol. Lett. 162:75-82(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4001;
RA Fudyk T.C., Maclean I.W., Simonsen J.N., Njagi E.N., Kimani J.,
RA Brunham R.C., Plummer F.A.;
RT "Genetic Diversity and Mosaicism at the por locus of Neisseria
RT gonorrhoeae."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
DR EMBL: AF015118; AAC38564.1; -.
DR EMBL: AF090823; AAC79492.1; -.
DR InterPro: IPR001702; Gram_neg_porin.
DR Pfam: PF00267; Gram-ve_porins; 1.
DR PRINTS: PR00182; ECOLNEIPORIN.
DR PROSITE: PS00576; GRAM_NEG_PORIN; 1.
DR Outer membrane; Porin; Signal; Transmembrane.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 326 OUTER MEMBRANE PROTEIN I.
FT SEQUENCE 326 AA; 35577 MW; BF50F03EAF73BEE4 CRC64;
SQ

Query Match 35.9%; Score 52; DB 2; Length 326;
Best Local Similarity 43.5%; Pred. No. 6.2;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEPN 23
| | | | | : : : | | | : :
Db 268 KGLVYHADLSNDYDQVYVGAEPD 290

RESULT 6
Q924L5 PRELIMINARY; PRT; 327 AA.
AC Q924L5;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, last annotation update)
DE OUTER MEMBRANE PROTEIN I PRECURSOR.
GN POR.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4266;
RA Fudyk T.C., Maclean I.W., Simonsen J.N., Njagi E.N., Kimani J.,
RA Brunham R.C., Plummer F.A.;
RT "Genetic Diversity and Mosaicism at the por locus of Neisseria
RT gonorrhoeae."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
DR EMBL: AF090808; AAC79477.1; -.
DR InterPro: IPR001702; Gram_neg_porin.

```

```

DR Pfam: PF00267; Gram-ve_porins; 1.
DR PRINTS: PR00182; ECOLNEIPORIN.
DR PROSITE: PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Porin; Signal; Transmembrane.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 327 OUTER MEMBRANE PROTEIN I.
FT SEQUENCE 327 AA; 35471 MW; FDB0D0FE8E53E32F CRC64;
SQ

Query Match 35.9%; Score 52; DB 2; Length 327;
Best Local Similarity 43.5%; Pred. No. 6.3;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEPN 23
| | | | | : : : | | | : :
Db 269 KGLVYHADLSNDYDQVYVGAEPD 291

RESULT 7
O13834 PRELIMINARY; PRT; 3227 AA.
AC O13834;
DT 01-JUN-1998 (TReMBLrel. 05, Created)
DT 01-JUN-1998 (TReMBLrel. 05, last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, last annotation update)
DE UBICUTIN SYSTEM PROTEIN.
GN SPAC19D5.04.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Devlin K., Churcher C.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Bartell B.G., Rajandream M.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z99531; CAB16714.1; -.
DR InterPro: IPR000569; HECT.
DR InterPro: IPR001000; Glyco_hydro_10.
DR InterPro: IPR001345; PG_mutase.
DR Pfam: PF00632; HECT; 1.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; UNKNOWN_1.
DR PROSITE: PS50237; HECT; 1.
DR PROSITE: PS00175; PG_MUTASE; UNKNOWN_1.
DR SMART: SM00119; HECTC; 1.
SQ SEQUENCE 3227 AA; 365031 MW; 07FC47AB79124575 CMC64;

Query Match 35.9%; Score 52; DB 3; Length 3227;
Best Local Similarity 46.2%; Pred. No. 80;
Matches 12; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CIGYHLDVSLAFSEISVGAEPNKDDC 27
| : | : | | | : | | | : | |
Db 1339 CILYLLLEVLADSKKPDPEFFNSDC 1364

RESULT 8
O917L2 PRELIMINARY; PRT; 241 AA.
AC O917L2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, last annotation update)
DE CG3066 PROTEIN.
GN CG3066.
OS Drosophila melanogaster (Fruit fly).

```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Ephydrophoridae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA April J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu P.V., Bertan B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Harris N.L., Harvey D., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iregyan C.,  
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.R., Houck J.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradlin A.C., Stapleton M., Strong R., Wang A.H., Wang X.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zheng F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.  
 CC EMBL: AEO03676; AAC22127.1; -  
 DR EMBL: AEO03676; AAC22127.1; -  
 DR FlyBase: FBgn0037515; CG3066.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR001254; Trypsin.  
 DR Pfam: PF00089; Trypsin.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00020; Tryp-Spc. 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS. UNKNOWN\_1.  
 DR PROSITE: PS00135; TRYPSIN\_SER. 1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 241 AA; 27002 MW; AAA0228330F64C1E CRC64;

Query Match 32.4%; Score 47; DB 5; Length 241;  
 Best Local Similarity 40.0%; Pred. No 27;  
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 8 DVSLAFSEISVGAEFNKDDC 27  
 Db 168 NIHLSQLCYGGEFFRDC 187

RESULT 9  
 O9SSX1 PRELIMINARY; PRT; 278 AA.

AC O9SSX1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE OSEYAL.  
 GN OSEYAL.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzaeae; Oryza.  
 ON NCBI\_TaxID=4530;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99430869; PubMed=10503544;  
 RA Takeda Y., Hatanoto S., Sentoku N., Matsuno M.;  
 RT "Homologs of animal eyes absent (eya) genes are found in higher  
 plants."  
 RL Mol. Gen. Genet. 262:131-138(1999).  
 RL EMBL: AB028887; BAA5161.1;  
 SQ SEQUENCE 278 AA; 31044 MW; 7BC1E2C41343B7A1 CRC64;

Query Match 32.4%; Score 47; DB 10; Length 278;  
 Best Local Similarity 40.7%; Pred. No 31;  
 Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
 Db 202 KCLLYRDLDDVAFENYSSWVGKLC 228

RESULT 10  
 O9SSNG2 PRELIMINARY; PRT; 367 AA.  
 AC O9SSNG2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE EST C72179(E1153) CORRESPONDS TO A REGION OF THE PREDICTED GENE.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzaeae; Oryza.  
 ON NCBI\_TaxID=4530;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC  
 clone: P0514G12."  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000616; BAA5447.1; -  
 SQ SEQUENCE 367 AA; 42367 MW; FF9F290DE82F273E CRC64;

Query Match 32.4%; Score 47; DB 10; Length 367;  
 Best Local Similarity 40.7%; Pred. No. 42;  
 Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

OY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
 Db 235 KCLLYRDLDDVAFENYSSWVGKLC 261

RESULT 11  
 O9V322 PRELIMINARY; PRT; 391 AA.  
 ID O9V322;  
 AC O9V322;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CG3066 PROTEIN.  
 GN CG3066.



OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephyraoidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP STRAIN=BERKELEY;  
RC MEDLINE=20196006; PubMed=10731132;  
RX  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mkicos G.L.G.,  
RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailes R.M., Bass A.V., Baxendale J., Bayraktoglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck Y., Brokstein P., Brotlier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwu C.,  
RA Jaitani M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weisenick G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of Drosophila melanogaster.",  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RP Ryu J.H., Lee W.J.;  
RT "Molecular cloning of a Drosophila serine proteinase homologous to  
RT easter.",  
RT Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
RL  
CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.  
CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).  
DR EMBL: AE003676; AAF54143.1; -  
DR EMBL: AF233093; AAF43410.1; -  
DR HSSP: P20160; 1A75  
DR FlyBase: FBgn0037515; CG3066.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Trypsin.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; TRYP\_SPC\_1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SSR; 1.  
DR HydroLase: Serine protease.  
SQ SEQUENCE 391 AA; 43127 MW; 3E1416F12DDF89F6 CRC64;

|        |   |                      |                         |
|--------|---|----------------------|-------------------------|
| OY     | 8   | DVSLAFSEISVGAEFNKDDC | 27                      |
|        | ::   ::         :   |                      |                         |
| Db     | 318   | NHLLSSQLCVGGEFYRDSC  | 337                     |
| <br>   |   |                      |                         |
| RESULT | 12  |                      |                         |
| ID     | 045247  | PRELIMINARY;         | PRT; 1002 AA.           |
| OC     | 045247:   |                      |                         |
| DT     | 01-JUN-1998 (TEMBrel. 06,   |                      | Created)                |
| DT     | 01-MAY-2000 (TEMBrel. 13,   |                      | Last sequence update)   |
| DT     | 01-MAR-2001 (TEMBrel. 16,   |                      | Last annotation update) |
| DE     | C13B4.1 PROTEIN.  |                      |                         |
| GN     | C13B4.1.  |                      |                         |
| OS     | Caenorhabditis elegans.   |                      |                         |
| OC     | Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea |                      |                         |
| OC     | Rhabdilitidae; Peioderinae; Caenorhabditis.                           |                      |                         |
| OX     | NCBI_TaxID=6239;  |                      |                         |
|        | [1]   |                      |                         |
| RN     | SEQUENCE FROM N.A.  |                      |                         |
| RA     | Smye R.;  |                      |                         |
| RL     | Submitted (NOV-1996) to the EMBL/genbank/DBJ databases.               |                      |                         |
| RP     | [2]   |                      |                         |
| RP     | SEQUENCE FROM N.A.  |                      |                         |
| RX     | MEDLINE=99069613; PubMed=9851916;                                     |                      |                         |
| RA     | none;   |                      |                         |
| RT     | "Genome sequence of the nematode C.elegans: A platform for            |                      |                         |
| RT     | investigating biology."   |                      |                         |
| RL     | Science 282:2012-2018(1998).  |                      |                         |
| DR     | EMBL; Z81468; CAB03875.2.-.   |                      |                         |
| SO     | SEQUENCE  |                      |                         |
|        | 1002 AA; 109008 MW; 01AEA804ED30FAE CRC64;                            |                      |                         |

```

QY      1 KCLGYHLDVS-----LAFSEISVGAEFNKDDC 27
          | | | | | | | | | | | | | | | | | | |
Db      170 KGLGYHKDQSSVSRRKRLRLESSSPG--PQSSQC 199
          | | | | | | | | | | | | | | | | | | |
Query Match 59.21%   Score 46.5; DB 5; Length 1002;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 14; Conservative 3; Mismatches 8; Indels 7; Gaps 2.

RESULT  13
Q9YMI5      PRELIMINARY;      PRT;      203 AA.
Q9YMI5
AC  Q9YMI5;
DT  01-MAY-1999 (TREMBlrel. 10, Created)
DT  01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT  01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE  LIDRF-143 PEPTIDE.
DR  LMANTRIA dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
OS  Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC  Nucleopolyhedrovirus.
CX  NCBI_TaxID=10449;
OX  NCBI_TaxID=10449;
RN  (1)
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99124785; PubMed=9887315;
RA  Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA  Slavicek J.M., Rohmann G.F.,
RT  "Sequence and analysis of the genome of a baculovirus pathogenic for
RL  Lymantria dispar.";
RN  Virology 253:17-34(1999).
RP  (2)
RP  SEQUENCE FROM N.A.
RA  Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA  Slavicek J., Rohmann G.F.;
RL  Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN  (3)
RP  SEQUENCE FROM N.A.
RA  Kuzio J.;
RL  Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
OR  EMBL; AF081810; AAC70329.1; -.

```

SQ SEQUENCE

203 AA; 22177 MW; 9D6E8BFC22DCBF73 CRC64;

## Query Match

Best Local Similarity 31.7%; Score 46; DB 12; Length 203;  
Matches 10; Conservative 4; Mismatches 6; Indels 6; Gaps 1;

QY 2 CIGYHLDVSLAFSEISVGAFFNKDDC 27  
 DB 143 CECGEJLDIDL-----VGRQFSANDC 162

## RESULT 14

09XEFO

ID 09XEFO PRELIMINARY; PRT: 254 AA.

AC 09XEFO: 01-NOV-1999 (TREMblrel. 12, Created)

DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)

DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)

GN HYPOTHEICAL 28.4 KDA PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicot; Rosidae;

OC NCBI\_TaxID=3702; Brassicales; Brassicaceae; Arabidopsis.

[1]

SEQUENCE FROM N.A.

RX MEDLINE=99225673; PubMed=10207155;

RA Wang M.L., Belmonte S., Kim U., Dolan M., Morris J.W., Goodman H.M.,

RT 707M07.\*;

RL Genome Res. 9:325-333(1999).

[2]

SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

RA Adams M.D., Shenn M., Vanden S.E., Umayan L., Tallon L.J., Gill J.E.,

RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,

RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";

RL Nature 402:761-768(1999).

DR EMBL; AF085279; AAD25935.1; -

DR EMBL; AC018721; AAF18734.1; -

DR InterPro; IPR001092; HLH\_dim.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

KW Hypothetical protein.

SQ SEQUENCE 254 AA; 28414 MW; EE3BCE353EE52619 CRC64;

## Query Match

Best Local Similarity 31.7%; Score 46; DB 10; Length 254;  
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 SLAFSEISVGAFFNKDDC 27  
 DB 36 NLGFSFSGFNFPADDC 53

## RESULT 15

067397

AC 067397 PRELIMINARY; PRT: 292 AA.

DT 01-AUG-1998 (TREMblrel. 07, Created)

DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

DE 01-MAY-2000 (TREMblrel. 13, Last annotation update)

GN HYPOTHEICAL 33.2 KDA PROTEIN.

AO\_1392.

OS Aquifex aeolicus.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.

OX NCBI\_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,

RT "The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus.";

RL Nature 392:353-358(1998).

DR EMBL; AE000738; AAC07368.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 292 AA; 33244 MW; 38AB3FD436D35FEB CRC64;

## Query Match

Best Local Similarity 31.7%; Score 46; DB 2; Length 292;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 IGYHLDVSLAFSEISVG 19  
 DB 261 LKYEFLRLAFSATLTG 277

Search completed: January 6, 2002, 09:10:38  
 Job time: 451 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:01:52 ; Search time 42.93 Seconds

(Without alignments)  
46.587 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                 |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1          | 145   | 100.0       | 27     | 20    | AAV27328 Human C9 protein f |
| 2          | 142   | 97.9        | 82     | 20    | AAV27324 Human C9 protein f |
| 3          | 142   | 97.9        | 560    | 18    | AAV18310 Human complement C |
| 4          | 57    | 39.3        | 86     | 20    | AAI27325 Rabbit C9 protein  |
| 5          | 57    | 39.3        | 561    | 18    | AAV18311 Rabbit complement  |
| 6          | 50    | 34.5        | 1263   | 22    | AAV72637 Exophiala spinifer |
| 7          | 44.5  | 30.7        | 162    | 21    | AAV32113 Mouse 15 kDa selen |
| 8          | 44    | 30.3        | 582    | 22    | AAU02929 Angiotensin conver |
| 9          | 44    | 30.3        | 662    | 21    | AA808632 Amino acid sequenc |
| 10         | 44    | 30.3        | 1511   | 21    | AA638528 Arabidopsis thalia |
| 11         | 44    | 30.3        | 1544   | 21    | AA638527 Arabidopsis thalia |

|    |      |      |      |    |                             |
|----|------|------|------|----|-----------------------------|
| 12 | 44   | 30.3 | 1547 | 21 | AA638526 Arabidopsis thalia |
| 13 | 43.5 | 30.0 | 330  | 22 | AA662369 Follicular conjunc |
| 14 | 43   | 29.7 | 82   | 21 | AA654363 Human pancreatic c |
| 15 | 43   | 29.7 | 82   | 22 | AA674079 Human colon cancer |
| 16 | 43   | 29.7 | 90   | 21 | AA641682 Human ORFX ORF1446 |
| 17 | 43   | 29.7 | 154  | 21 | AA638111 Arabidopsis thalia |
| 18 | 43   | 29.7 | 168  | 21 | AA638110 Arabidopsis thalia |
| 19 | 43   | 29.7 | 213  | 20 | AAV27646 Human secreted pro |
| 20 | 43   | 29.7 | 271  | 21 | AA631287 Arabidopsis thalia |
| 21 | 43   | 29.7 | 282  | 21 | AA631664 Arabidopsis thalia |
| 22 | 43   | 29.7 | 282  | 21 | AA640090 Arabidopsis thalia |
| 23 | 43   | 29.7 | 285  | 21 | AA631286 Arabidopsis thalia |
| 24 | 43   | 29.7 | 306  | 21 | AA613663 Arabidopsis thalia |
| 25 | 43   | 29.7 | 306  | 21 | AA640089 Arabidopsis thalia |
| 26 | 43   | 29.7 | 309  | 21 | AA631285 Arabidopsis thalia |
| 27 | 43   | 29.7 | 320  | 21 | AA613662 Arabidopsis thalia |
| 28 | 43   | 29.7 | 320  | 21 | AA640088 Arabidopsis thalia |
| 29 | 43   | 29.7 | 567  | 18 | AAW20055 Saccharomyces ope  |
| 30 | 43   | 29.7 | 1070 | 21 | AA61124 Human P13 kinase p  |
| 31 | 42.5 | 29.3 | 434  | 22 | AA682564 S. epidermidis ope |
| 32 | 42.5 | 29.3 | 591  | 16 | AA674802 Arabidopsis thalia |
| 33 | 42   | 29.0 | 58   | 21 | AA646705 Arabidopsis thalia |
| 34 | 42   | 29.0 | 74   | 21 | AA646704 Arabidopsis thalia |
| 35 | 42   | 29.0 | 75   | 21 | AA646703 Arabidopsis thalia |
| 36 | 42   | 29.0 | 77   | 22 | AA675197 Human colon cancer |
| 37 | 42   | 29.0 | 117  | 19 | AAW51102 Flag-amyloid prote |
| 38 | 42   | 29.0 | 148  | 19 | AAW56138 Open reading frame |
| 39 | 42   | 29.0 | 298  | 22 | AA691621 S. epidermidis ope |
| 40 | 42   | 29.0 | 305  | 22 | AA682721 Maize adenylosucci |
| 41 | 42   | 29.0 | 484  | 17 | AA67734 Maize adenylosucci  |
| 42 | 42   | 29.0 | 484  | 17 | AAW02615 Maize adenylosucci |
| 43 | 42   | 29.0 | 484  | 20 | AAW94455 Arabidopsis thalia |
| 44 | 42   | 29.0 | 516  | 21 | AA646683 Arabidopsis thalia |
| 45 | 42   | 29.0 | 544  | 21 | AA646682 Arabidopsis thalia |

#### ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 | AAV27328 standard; peptide: 27 AA.                                   |
| ID       | AAV27328   |
| XX       | AAV27328:  |
| AC       | 05-NOV-1999 (first entry)  |
| XX       |  |
| DF       | Human C9 protein fragment (residues 359-384).                        |
| XX       |  |
| DE       | CD59 mediated complement; human; Cd59 protein; mimetic;              |
| XX       | tumour therapy; complement-mediated inflammation; immune disorder;   |
| KW       | immunovascultitis; rheumatoid arthritis; scleroderma; C5b-9 complex. |
| KW       |  |
| XX       |  |
| OS       | Homo sapiens.  |
| XX       |  |
| PN       | MO9940115-A2.  |
| XX       |  |
| PD       | 12-AUG-1999.   |
| XX       |  |
| PF       | 09-FEB-1999; 99WO-0502802.   |
| XX       |  |
| PR       | 09-FEB-1998; 98US-0020393.   |
| XX       |  |
| PA       | (BLOO-) BLOOD CENT RES FOUND INC.                                    |
| PA       | (OKLA-) OKLAHOMA MEDICAL RES FOUND.                                  |
| XX       |  |
| PI       | Sims PJ;   |
| XX       |  |
| DR       | WPI; 1999-527301/44.   |
| XX       |  |
| PT       | Compounds modulating CD59 mediated complement activity, for          |
| XX       | treatment of, e.g. Immunovascultitis                                 |
| XX       |  |

PS Example 2; Page 36; 75pp; English.

CC The invention relates to compounds modulating CD59 mediated complement  
CC activity. It provides (i) molecules structurally mimicking human CD59  
CC amino acid residues 42-58 (region which serves as binding site for CD59  
CC -C9 interactions) when they are in a spatial orientation which can  
CC inhibit the formation of the human C5b-9 complex. These mimetics  
CC specifically bind to amino acid residues 359-384 of human C9. (11)  
CC In a spatial orientation which can promote the formation of the C5b-9  
CC complex. Compounds that mimic CD59 can be used to increase CD59  
CC inhibition of C5b-9 complex assembly. This is especially useful in  
CC patients in need of suppression of complement-mediated inflammation, e.g.  
CC immune disorders and diseases such as immunovascularitis, rheumatoid  
CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote  
CC C5b-9 complex assembly. This is useful in patients in need of complement  
CC activation. The composition can be administered as an adjunct to tumour  
CC therapy. The present sequence represents a human C9 protein fragment.

XX Sequence 27 AA;

Query Match 100.0%; Score 145; DB 20; Length 27;  
Best Local Similarity 100.0%; Pred. No. 3 4e-17;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
DB 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

RESULT 2  
AAV27324  
ID AAV27324 standard; peptide; 82 AA.

AC AAV27324;

DT 05-NOV-1999 (first entry)

DE Human C9 protein fragment (residues 334-415).

KW CD59 mediated complement; human; CD59 protein; C9 protein; mimetic;  
KW tumour therapy; complement-mediated inflammation; immune disorder;  
KW immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex.

OS Homo sapiens.

PN MO9940115-A2.

PD 12-AUG-1999.

PF 09-FEB-1999; 99WO-US02802.

PR 09-FEB-1998; 98US-0020393.

PA (BLOO-) BLOOD CENT RES FOUND INC.  
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

PI Sims PJ;

DR WPI; 1999-527301/44.

PT Compounds modulating CD59 mediated complement activity, for  
PT treatment of, e.g. immunovascularitis

PS Disclosure; Fig 10; 75pp; English.

CC The invention relates to compounds modulating CD59 mediated complement  
CC activity. It provides (i) molecules structurally mimicking human CD59  
CC amino acid residues 42-58 (region which serves as binding site for CD59  
CC -C9 interactions) when they are in a spatial orientation which can  
CC inhibit the formation of the human C5b-9 complex. These mimetics  
CC specifically bind to amino acid residues 359-384 of human C9. (11)

CC molecules structurally mimicking C9 amino acids 359-384 when they are  
CC in a spatial orientation which can promote the formation of the C5b-9  
CC complex. Compounds that mimic CD59 can be used to increase CD59  
CC inhibition of C5b-9 complex assembly. This is especially useful in  
CC patients in need of suppression of complement-mediated inflammation, e.g.  
CC immune disorders and diseases such as immunovascularitis, rheumatoid  
CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote  
CC C5b-9 complex assembly. This is useful in patients in need of complement  
CC activation. The composition can be administered as an adjunct to tumour  
CC therapy. The present sequence represents a human C9 protein fragment.

XX Sequence 82 AA;

Query Match 97.9%; Score 142; DB 20; Length 82;  
Best Local Similarity 96.3%; Pred. No. 4 2e-16;

Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
DB 25 KCLGYHLDVSLAFSEISVGAEFNKDDC 51

RESULT 3

AAW18310  
ID AAW18310 standard; Protein; 560 AA.

AC AAW18310;

DT 10-AUG-1997 (first entry)

DE Human complement C9.

KW Complement C9; CD59; C5b9 complex; tumour; therapy; contraceptive;  
KW antibody; inflammation.

OS Homo sapiens.

FN Key Location/Qualifiers  
FT Binding-site 356..437  
FT /label= CD59 binding-site  
FT /note= "Claim 1 (amino acids 359-384)"

FT Disulfide-bond 381..406

XX MO9717987-A1.

PD 22-MAY-1997.

PF 08-NOV-1996; 96WO-US17940.

PR 15-NOV-1995; 95US-0559492.

PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

PI Sims PJ;

DR WPI; 1997-289058/26.

DR N-PSDB; AAT68886.

PT Modulating binding of CD59 to C9 complement component - uses agent  
PT that mimics or binds the C9-specific motif, used to activate  
PT complement in tumour therapy or to treat complement mediated  
PT inflammation

PS Disclosure; Page 33-35; 51pp; English.

CC Amino acid residues 359-384 of human complement C9 are critical  
CC for binding CD59 to C9, resulting in inhibition of C5b-9 complex  
CC assembly. Peptides that mimic this human C9-derived peptide  
CC sequence, and antibodies raised against such peptides, can be used  
CC to modulate binding of CD59 to C9. The peptides bind to endogenous  
CC CD59 to prevent the CD59 from inhibiting assembly of the C5b-9  
CC complex, thereby increasing complement-mediated activation of cells.

CC Inhibition of CD59 is useful as an adjuvant for tumour therapy and  
 CC as a contraceptive. The antibodies inhibit assembly of the C5b-9  
 CC complex by binding to C9. This is useful for disorders  
 CC characterised by excessive complement activation or complement-  
 CC mediated cytotoxicity.

XX Sequence 560 AA;

Query Match 97.9%; Score 142; DB 18; Length 560;  
 Best Local Similarity 96.3%; Pred. No. 4.5e-15;  
 Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
 :|||||  
 Db 380 rclqghldvslafseisvgaefnkdcc 406

RESULT 4

AAV27325 AAV27325 standard; peptide; 86 AA.

XX AAV27325;

DT 05-NOV-1999 (first entry)

XX Rabbit C9 protein fragment.

DE CD59 mediated complement; human; CD59 protein; C9 protein; mimetic;

KW tumour therapy; complement-mediated inflammation; immune disorder;

KW Immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex;

KW rabbit.

XX Oryctolagus sp.

XX WO9940115-A2.

XX 12-AUG-1999.

PD 09-FEB-1999; 99WO-US02802.

XX 09-FEB-1998; 98US-0020393.

PA (BLOO-) BLOOD CENT RES FOUND INC.  
 PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX Sims PJ;

DR WPI; 1999-527301/44.

XX Compounds modulating CD59 mediated complement activity, for  
 PT treatment of, e.g. immunovascularitis

XX Disclosure; Fig 10; 75pp; English.

XX The invention relates to compounds modulating CD59 mediated complement  
 CC activity. It provides (1) molecules structurally mimicking human CD59  
 CC amino acid residues 42-58 (region which serves as binding site for C5b-9  
 CC -C9 interactions) when they are in a spatial orientation which can  
 CC inhibit the formation of the human C5b-9 complex. These mimetics  
 CC specifically bind to amino acid residues 359-384 of human C9. (11)  
 CC molecules structurally mimicking C9 amino acids 359-384 when they are  
 CC in a spatial orientation which can promote the formation of the C5b-9  
 CC complex. Compounds that mimic CD59 can be used to increase CD59  
 CC inhibition of C5b-9 complex assembly. This is especially useful in  
 CC patients in need of suppression of complement-mediated inflammation, e.g.  
 CC immune disorders and diseases such as immunovascularitis, rheumatoid  
 CC arthritis, scleroderma. Compounds that mimic C9 can be used to promote  
 CC C5b-9 complex assembly. This is useful in patients in need of complement  
 CC activation. The composition can be administered as an adjunct to tumour  
 CC therapy. The present sequence represents a rabbit C9 protein fragment.

XX Sequence 86 AA;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
 :|||||  
 Db 388 rclqghldvslafseisvgaefnkdcc 418

Query Match 39.3%; Score 57; DB 20; Length 86;  
 Best Local Similarity 38.7%; Pred. No. 0.063;  
 Matches 12; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

RESULT 5  
 AAW18311 AAW18311 standard; Protein; 561 AA.  
 XX AAW18311;  
 AC AAW18311;  
 XX 10-AUG-1997 (first entry)  
 DT Rabbit complement C9.  
 DE Rabbit complement C9.  
 XX Complement C9; CD59; C5b9 complex; tumour; therapy; contraceptive;  
 KW antibody; inflammation.  
 OS Oryctolagus cuniculus.  
 XX Key Location/Qualifiers  
 FT Disulfide-bond 389..418  
 PN WO9717987-A1.  
 XX 22-MAY-1997.  
 PD 08-NOV-1996; 96WO-US17940.  
 XX 15-NOV-1995; 95US-0559492.  
 XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
 PA Sims PJ;  
 PI WPI; 1997-289058/26.  
 DR N-PSDB; T688887.  
 XX Modulating binding of CD59 to C9 complement component - uses agent  
 PT that mimics or binds the C9-specific motif, used to activate  
 PT complement in tumour therapy or to treat complement mediated  
 PT inflammation  
 XX Disclosure; Page 36-38; 51pp; English.

XX A polypeptide (AAW18311) comprises rabbit complement C9. Chimeric  
 CC proteins were constructed in which the segment of C9 corresponding  
 CC to the putative CD59 binding site were interchanged between rabbit  
 CC and human C9 (see also AAW18310). These were tested for haemolytic  
 CC activity and for sensitivity to inhibition by membrane CD59.  
 CC Substr. of human C9 residues 334-415 into rabbit C9 resulted in a  
 CC protein that was indistinguishable from human C9 in its sensitivity  
 CC to inhibition by CD59. When this same segment of human C9 was  
 CC replaced by the corresponding rabbit sequence (aa339-424), the  
 CC chimera was indistinguishable from rabbit C9 and was virtually  
 CC unaffected by the presence of membrane CD59.

XX Sequence 561 AA;

RESULT 6  
 ID AAY72637 standard; Protein: 1263 AA.  
 AC AAY72637;  
 DT 02-MAY-2001 (first entry)  
 DE Exophiala splnifera p-glycoprotein, a fumonisin degradative enzyme.  
 KW Fumonisin; degradative enzyme; transport enzyme; flavin monooxygenase;  
 KW aldehyde dehydrogenase; permease; p-glycoprotein; fumonisin esterase;  
 KW detoxification; mycotoxin; animal feed; human feed; silage;  
 KW transgenic plant; transgenic animal; microbial spray.  
 OS Exophiala splnifera.  
 FH Key Location/Qualifiers  
 FT MISC-difference 157 /label= Unknown  
 FT /note= "Encoded by CAR"  
 PN W0200105980-A1.  
 PD 25-JAN-2001.  
 PF 14-JUL-1999; 99WO-US15824.  
 PR 14-JUL-1999; 99WO-US15824.  
 PA (PION-) PIONEER HI-BRED INT. INC.  
 PA (CURA-) CORAGEN CORP.  
 PI Duvick JP, Maddox J, Gilliam J, Folkerts O, Crasta OR;  
 DR WPI: 2001-147345/15.  
 DR N-PSDB; AAD02694, AAD02695.  
 PT Novel polynucleotides encoding Exophiala degradative or transport  
 PT enzyme which is useful for detoxifying fumonisin or structurally  
 PT related mycotoxin during processing of grain for human or animal food  
 PS consumption -  
 PS  
 PS Claim 1c; Page 77-82; 90pp; English.  
 CC The patent discloses novel polynucleotides encoding Exophiala splnifera  
 CC fumonisin degradative or transport enzymes such as flavin monooxygenase,  
 CC aldehyde dehydrogenase, permease and p-glycoprotein. A fumonisin esterase  
 CC enzyme with at least one fumonisin degradative enzyme is useful for  
 CC detoxifying fumonisin or a structurally related mycotoxin present in  
 CC harvested grain, on application to a plant or to harvested grain  
 CC during processing, or to processed grain that is to be used as animal  
 CC or human feed, or as silage. Nucleotide fragments of the present  
 CC invention are useful as probes and primers. They can be introduced  
 CC into microorganisms that multiply on plants to deliver enzymes to  
 CC potential target crops. The genes encoding the degrading enzymes to  
 CC is supplied to a vector into a microbial host and the transformed host  
 CC pathogenicity of a fungus producing fumonisin. The genes of the  
 CC invention are fermented in a bacterial host and the resulting bacteria  
 CC can be used alone or in combination to engineer microbes or other  
 CC organisms to metabolise fumonisin and resist its toxic effects.  
 CC The present protein sequence is p-glycoprotein, a fumonisin  
 CC degradative enzyme from Exophiala splnifera.  
 SQ Sequence 1263 AA:

Query Match 34.5%; Score 50; DB 22; Length 1263;  
 Best Local Similarity 42.1%; Pred. No. 25;

Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KCLGYHLDVSLAFSEISVG 19  
 Db 922 rcfgfhlsqsmeflatalg 940  
 RESULT 7  
 ID AAY32113 standard; Protein: 162 AA.  
 AC AAY32113;  
 DT 01-FEB-2000 (first entry)  
 DE Mouse 15 kDa selenoprotein.  
 KW Selenoprotein; selenium; differential expression; tumour; mouse;  
 KW prostate cancer; diagnosis; polymorphism.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT Modified-site 93 /label= OTHER  
 FT /note= "selenocysteine"  
 PN W09951637-A1.  
 PD 14-OCT-1999.  
 PF 06-APR-1999; 99WO-US07560.  
 PR 06-APR-1998; 98US-0080850.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA Gladyshev VN, Wootton JC, Hatfield DL, Jeang K;  
 DR WPI: 2000-013045/01.  
 DR N-PSDB; AA34469.  
 PT New isolated selenoprotein polypeptides, used to develop products for  
 PT detecting susceptibility to or treating cancers e.g. prostate cancers  
 PS -  
 PS Claim 1; Page 60; 67pp; English.  
 CC The present sequence represents a mouse 15 kDa selenoprotein that  
 CC is differentially expressed in cancer cells, such as cancer cells.  
 CC It includes a selenocysteine residue at position 93 that is encoded  
 CC by a TGA codon in the cDNA (see AA34468). Polymorphisms in the  
 CC human 15 kDa selenoprotein gene are associated with susceptibility  
 CC to cancer. Claimed methods of detecting the presence, or of  
 CC quantifying the level of expression, of the 15 kDa selenoprotein  
 CC involve contacting a sample with a specific binding agent, such as  
 CC a polyclonal antibody or monoclonal antibody, that specifically  
 CC binds to the 15 kDa selenoprotein. A claimed method for dietary  
 CC mammalian 15 kDa selenoprotein. A claimed method for dietary  
 CC below normal, enhancing the level by providing additional selenium  
 CC in the diet. The susceptibility to cancer of a subject having an  
 CC increased predetermined genetic susceptibility to cancer is reduced  
 CC by administering a mammalian 15 kDa selenoprotein or by  
 CC overexpressing the selenoprotein using gene therapy.  
 SQ Sequence 162 AA:

Query Match 30.7%; Score 44.5; DB 21; Length 162;  
 Best Local Similarity 42.3%; Pred. No. 16;  
 Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

OY 3 LGYHLDVSLAFSEIS-VGAEPNKKDDC 27  
 11 : : 11 : : 1111 : 1  
 Db 13 Lgtrlllatafagsaigaefaseac 38

## RESULT 8

AAU02929 standard; Protein: 582 AA.  
 AAU02929;

12-SEP-2001 (first entry)

Angiotensin converting enzyme (ACEV) splice variant protein #29.

Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
 granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
 platelet-derived endothelial cell growth factor; cardiovascular disease;  
 cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;  
 vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
 myocardial infarction; coronary arterial thrombosis; renal disease;  
 diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
 multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
 nonarctoidotic pulmonary granulomatous disease; endothelial abnormality;  
 vascular disorder; asbestosis.

Mus sp..

MO200136632-A2.

25-MAY-2001.

17-NOV-2000; 2000MO-IL00766.

17-NOV-1999; 99IL-0132978.  
 10-DEC-1999; 99IL-0134355.

(COMP-) COMPUGEN LTD.

Levine Z, David A, Azar I, Khosravi R, Bernstein J;

WPI; 2001-336004/35.

N-PSDB; AAS06029.

Novel alternative splicing variants e.g. variant of angiotensin

converting enzyme (ACEV), useful in identifying candidate compounds

capable of binding to the variant and to detect anti-variant antibodies

Claim 4; Fig 29; 519pp; English.

The sequence represents an angiotensin converting enzyme splice variant of (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonarctoidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis.

Sequence 582 AA;

Query Match 30.3%; Score 44; DB 22; Length 582;  
 Best Local Similarity 47.8%; Pred. No. 97;

Matches 11; Conservative 4; Mismatches 6; Indels 2; Gaps 1;  
 OY 4 GYHLDVSLAFSEISVGAEPNKKDD 26  
 11 : 1 : 1111 1 : 1 : 1  
 Db 490 gynlspslpiseis--aqcnrgd 510

## RESULT 9

AA08632 standard; Protein: 662 AA.

20-DEC-2000 (first entry)

Amino acid sequence of a Staphylococcus aureus tktA polypeptide.

tktA; microbial infection; microbial vaccine; antibacterial.  
 Staphylococcus aureus.

MO200050435-A1.

31-AUG-2000.

18-FEB-2000; 2000MO-US04444.

25-FEB-1999; 99US-0257438.

(SMIK) SMITHKLINE BEECHAM CORP.

McDevitt D, Yu J, Lonetto MA, Jaworski DD, Wang M, Warren RL;

Li T;

WPI; 2000-533177/48.

N-PSDB; AAA64393.

Isolated tktA polypeptide used for treating microbial infections and detecting agonists and antagonists to the polypeptide -

Claim 1; Page 4; 41pp; English.

The present sequence represents a tktA polypeptide from Staphylococcus aureus. The polypeptides and polynucleotides are useful for treatment of microbial infections, identifying agonists and antagonists for treating microbial infections and conditions associated with such infections and in diagnostic assays for detecting disease associated with microbial infections. The polypeptides and polynucleotides are also useful as microbial vaccines. The polypeptides are also used to screen for antibacterial compounds.

Sequence 662 AA;

Query Match 30.3%; Score 44; DB 21; Length 662;

Best Local Similarity 36.4%; Pred. No. 1.1e+02;

Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 4 GYHLDVSLAFSEISVGAEPNKKDD 25

Db 120 gfmavglataedhlagkfnke 141

## RESULT 10

AA038528 standard; Protein: 1511 AA.

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 47543.

Query Match 30.3%; Score 44; DB 22; Length 582;  
 Best Local Similarity 47.8%; Pred. No. 97;





PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161933.  
PR 29-OCT-1999; 99US-0162142.

Query Match 30.38; Score 44; DB 21; Length 1511;  
Best Local Similarity 47.68; Pred. No. 3.1e+02;  
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 6 HLDVSLAFSEISVGAFFNKD 26  
||||||| : :| :| :|  
Db 906 hldvslavawamqilfred 926

## RESULT 11

ID AAG38527 standard; Protein; 1544 AA.

AC AAG38527;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47542.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140891.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.

|    |              |               |
|----|--------------|---------------|
| PR | 22-SEP-1999; | 99US-0155139. |
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |
| PR | 12-OCT-1999; | 99US-0158363. |
| PR | 13-OCT-1999; | 99US-0158293. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 14-OCT-1999; | 99US-0159295. |
| PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. |
| PR | 18-OCT-1999; | 99US-0159638. |
| PR | 21-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160744. |
| PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-OCT-1999; | 99US-0160768. |
| PR | 21-OCT-1999; | 99US-0160770. |
| PR | 21-OCT-1999; | 99US-0160814. |
| PR | 21-OCT-1999; | 99US-0160815. |
| PR | 22-OCT-1999; | 99US-0160980. |
| PR | 22-OCT-1999; | 99US-0160981. |
| PR | 22-OCT-1999; | 99US-0160982. |
| PR | 25-OCT-1999; | 99US-0161404. |
| PR | 25-OCT-1999; | 99US-0161405. |
| PR | 25-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. |
| PR | 26-OCT-1999; | 99US-0161820. |
| PR | 28-OCT-1999; | 99US-0161992. |
| PR | 28-OCT-1999; | 99US-0161993. |
| PR | 29-OCT-1999; | 99US-0162142. |

|                       |  |                                    |           |              |
|-----------------------|--|------------------------------------|-----------|--------------|
| Query Match           | 30.3%  | Score 44;                          | DB 21;    | Length 1544; |
| Best Local Similarity | 47.6%  | Pred. No. 3.2e+02;                 |           |              |
| Matches 10;           | Conservative 3;  | Mismatches 8;                      | Indels 0; | Gaps 0;      |
| Oy                    | 6  | HLDSVIAFSEISVGAFFNKDD              | 26        |              |
|                       |  | : : :                              |           |              |
| Db                    | 939  | hldsvlavawafmgqifresd              | 959       |              |
| RESULT 12             |  |                                    |           |              |
| AAG38526              |  |                                    |           |              |
| ID                    | AAG38526   | standard; Protein; 1547 AA.        |           |              |
| XX                    |  |                                    |           |              |
| AC                    | AAG38526;  |                                    |           |              |
| XX                    |  |                                    |           |              |
| DT                    | 18-OCT-2000  | (first entry)                      |           |              |
| XX                    |  |                                    |           |              |
| DE                    | Arabidopsis thaliana   | protein fragment SEQ ID NO: 47541. |           |              |
| XX                    |  |                                    |           |              |
| KW                    | Protein identification; signal transduction pathway; metabolic pathway;  |                                    |           |              |
| KW                    | hybridisation assay; genetic mapping; gene expression control; promoter; |                                    |           |              |
| XX                    | termination sequence.  |                                    |           |              |
| OS                    | Arabidopsis thaliana.  |                                    |           |              |
| XX                    |  |                                    |           |              |
| FN                    | EP1033405-A2.  |                                    |           |              |
| PN                    |  |                                    |           |              |
| XX                    |  |                                    |           |              |
| PD                    | 06-SEP-2000.   |                                    |           |              |
| XX                    |  |                                    |           |              |
| PF                    | 25-FEB-2000; 2000EP-0301439.   |                                    |           |              |
| XX                    |  |                                    |           |              |
| PR                    | 25-FEB-1999; 99US-0121825.   |                                    |           |              |
| XX                    | 05-MAR-1999; 99US-0123180.   |                                    |           |              |

PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 04-MAY-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.

|        |   |
|--------|---|
| RESULT | 13  |
| ID     | AAG62569  |
| AC     | AAG62569 standard; Protein; 330 AA.                                 |
| DT     | 05-SEP-2001 (first entry)   |
| DE     | Follicular conjunctivitis related adenoviral protein SEQ ID NO: 10. |
| OS     | Mastadenovirus.   |
| PN     | JP2001095583-A.   |
| PD     | 10-APR-2001.  |
| PF     | 30-SEP-1999; 99JP-0278661.  |
| PR     | 30-SEP-1999; 99JP-0278661.  |
| PA     | (ITON/) ITO N.  |
| WI     | WI: 2001-341249/36.   |
| DR     | N-PSDB; AAA42033.   |
| PT     | New adenovirus for the prevention and treatment of Ad infection -   |
| XX     | Disclosure; Page 25-25; 45pp; Japanese.                             |

nm30773 to AAG9221 encode the human pancreatic cancer associated proteins, called pancreatic cancer antigens, given in AAB54008 to AAB54466. The human pancreatic cancer antigens have cytostatic, neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive, gynaecological, cardiant and antiinflammatory activities, and can be used in gene therapy. The polynucleotide and proteins can be used for preventing, treating, or ameliorating a medical condition or in assays for diagnosing a pathological condition or a susceptibility to one in a subject. Binding partners to the proteins and the activity of the proteins can be identified. The pancreatic cancer antigens can be used to detect, treat or prevent pancreatic disorders, especially cancer. Agonists and antagonists to the antigens can be screened for. The



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:01:52 ; Search time 25.89 Seconds  
(without alignments)  
23.468 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfillsl.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 142   | 97.9        | 82     | 2  | US-08-559-492-3   |
| 2          | 142   | 97.9        | 557    | 1  | US-08-313-288b-16 |
| 3          | 142   | 97.9        | 560    | 2  | US-08-559-492-5   |
| 4          | 57    | 39.3        | 86     | 2  | US-08-559-492-4   |
| 5          | 57    | 39.3        | 561    | 2  | US-08-559-492-12  |
| 6          | 42    | 29.0        | 117    | 2  | US-08-729-345-3   |
| 7          | 42    | 29.0        | 148    | 1  | US-08-565-386-8   |
| 8          | 42    | 29.0        | 359    | 4  | US-09-347-798-12  |
| 9          | 42    | 29.0        | 484    | 1  | US-08-361-611-4   |
| 10         | 42    | 29.0        | 484    | 1  | US-08-565-655-4   |
| 11         | 42    | 29.0        | 484    | 2  | US-08-946-967-4   |
| 12         | 41.5  | 28.6        | 365    | 3  | US-09-231-529-3   |
| 13         | 41.5  | 28.6        | 365    | 4  | US-08-977-816-3   |
| 14         | 41.5  | 28.6        | 619    | 3  | US-08-262-220-6   |
| 15         | 41.5  | 28.6        | 619    | 3  | US-08-471-733-6   |
| 16         | 41.5  | 28.6        | 619    | 3  | US-08-468-878-6   |
| 17         | 41.5  | 28.6        | 619    | 4  | US-08-750-494-6   |
| 18         | 41.5  | 28.6        | 990    | 2  | US-08-392-625-20  |
| 19         | 41.5  | 28.6        | 990    | 2  | US-08-466-961a-20 |
| 20         | 41.5  | 28.6        | 990    | 2  | US-08-645-193b-15 |
| 21         | 41    | 28.3        | 211    | 1  | US-08-631-607-4   |
| 22         | 41    | 28.3        | 560    | 2  | US-09-188-469-10  |
| 23         | 41    | 28.3        | 560    | 4  | US-09-397-238a-10 |
| 24         | 41    | 28.3        | 564    | 2  | US-08-948-569a-12 |
| 25         | 41    | 28.3        | 564    | 2  | US-09-188-469-12  |
| 26         | 41    | 28.3        | 564    | 4  | US-09-397-238a-12 |
| 27         | 41    | 28.3        | 564    | 4  | US-09-397-238a-12 |

|    |      |      |      |   |                   |                    |
|----|------|------|------|---|-------------------|--------------------|
| 28 | 41   | 28.3 | 2647 | 2 | US-08-583-562b-8  | Sequence 8, Appl1  |
| 29 | 41   | 28.3 | 2647 | 2 | US-08-779-113-8   | Sequence 8, Appl1  |
| 30 | 40   | 27.6 | 186  | 2 | US-08-960-022-18  | Sequence 18, Appl1 |
| 31 | 40   | 27.6 | 393  | 4 | US-09-347-798-18  | Sequence 18, Appl1 |
| 32 | 40   | 27.6 | 721  | 4 | US-09-146-221-4   | Sequence 4, Appl1  |
| 33 | 40   | 27.6 | 724  | 4 | US-09-146-221-6   | Sequence 6, Appl1  |
| 34 | 40   | 27.6 | 727  | 4 | US-09-146-221-8   | Sequence 8, Appl1  |
| 35 | 40   | 27.6 | 971  | 3 | US-09-112-450-2   | Sequence 2, Appl1  |
| 36 | 40   | 27.6 | 2471 | 3 | US-08-112-450-4   | Sequence 8, Appl1  |
| 37 | 39.5 | 27.2 | 621  | 3 | US-08-262-220-8   | Sequence 8, Appl1  |
| 38 | 39.5 | 27.2 | 621  | 3 | US-08-471-733-8   | Sequence 8, Appl1  |
| 39 | 39.5 | 27.2 | 621  | 3 | US-08-468-878-8   | Sequence 8, Appl1  |
| 40 | 39.5 | 27.2 | 621  | 4 | US-08-750-494-8   | Sequence 8, Appl1  |
| 41 | 39   | 26.9 | 282  | 3 | US-08-733-230-6   | Sequence 6, Appl1  |
| 42 | 39   | 26.9 | 282  | 4 | US-08-953-326-6   | Sequence 6, Appl1  |
| 43 | 39   | 26.9 | 529  | 1 | US-08-152-019a-40 | Sequence 40, Appl1 |
| 44 | 39   | 26.9 | 529  | 3 | US-08-650-599a-3  | Sequence 3, Appl1  |
| 45 | 39   | 26.9 | 603  | 3 | US-08-482-677-8   | Sequence 8, Appl1  |

#### ALIGNMENTS

```
RESULT 1
US-08-559-492-3
; Sequence 3, Application US/08559492
; Patent No. 5843884
;
GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-559-492-3
```

Query Match 97.9%; Score 142; DB 2; Length 82;  
Best local Similarity 96.3%; Pred. No. 1.8e-16;  
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27  
DB 25 RCLGYHLDVSLAFSEISVGAEFNKDDC 51

```
RESULT 2
US-08-313-288B-16
; Sequence 16, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jassell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NUMBER OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; US-08-313-288B-16

Query Match          97.9%; Score 142; DB 1; Length 557;
Best Local Similarity 96.3%; Pred. NO. 2.1e-15;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27
Db 377 RCLGYHLDVSLAFSEISVGAEFNKDDC 403

RESULT 3
US-08-559-492-5
; Sequence 5, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; US-08-559-492-5
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-559-492-5
```

```
Query Match          97.9%; Score 142; DB 2; Length 560;
Best Local Similarity 96.3%; Pred. NO. 2.1e-15;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27
Db 380 RCLGYHLDVSLAFSEISVGAEFNKDDC 406
```

```
RESULT 4
US-08-559-492-4
; Sequence 4, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-559-492-4
```





RESULT 9  
 US-08-361-611-4  
 ; Sequence 4, Application US/08361611  
 ; Patent No. 5519125  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Potter, Sharon L  
 ; APPLICANT: Ward, Eric R  
 ; TITLE OF INVENTION: Plant Adenylosuccinate Synthetase and  
 ; TITLE OF INVENTION: DNA Coding Therefor  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: CIBA-GEIGY Corporation  
 ; STREET: 7 Skyline Drive  
 ; CITY: Hawthorne  
 ; STATE: NY  
 ; COUNTRY: USA

```

1 RESULT 10
2 US-08-565-655-4
3 ; Sequence 4, Application US/08565655
4 ; Patent No. 5688939
5 ;
6 ; GENERAL INFORMATION:
7 ;
8 ; APPLICANT: Potter, Sharon L
9 ;
10 ; APPLICANT: Ward, Eric R
11 ;
12 ; TITLE OF INVENTION: Plant Adenylosuccinate Synthetase and
13 ; TITLE OF INVENTION: DNA Coding Therefor
14 ;
15 ; NUMBER OF SEQUENCES: 6
16 ;
17 ; CORRESPONDENCE ADDRESS:
18 ;
19 ; ADDRESSEE: Ciba Patent Department
20 ;
21 ; STREET: 540 White Plains Rd., POB 2005
22 ;
23 ; CITY: Tarrytown
24 ;
25 ; STATE: NY
26 ;
27 ; COUNTRY: USA
28 ;
29 ; ZIP: 10591-9005
30 ;
31 ; COMPUTER READABLE FORM:
32 ;
33 ; MEDIUM TYPE: Floppy disk
34 ;
35 ; COMPUTER: IBM PC compatible
36 ;
37 ; OPERATING SYSTEM: PC-DOS/MS-DOS
38 ;
39 ; SOFTWARE: PatentIn Release #1.0, Version #1.30B
40 ;
41 ; CURRENT APPLICATION DATA:
42 ;
43 ; APPLICATION NUMBER: US/08/565, 655
44 ;
45 ; FILING DATE:
46 ;
47 ; CLASSIFICATION: 210
48 ;
49 ; PRIOR APPLICATION DATA:
50 ;
51 ; APPLICATION NUMBER: US 08/361,611
52 ;
53 ; FILING DATE: 12-DEC-1994
54 ;
55 ; ATTORNEY/AGENT INFORMATION:
56 ;
57 ; NAME: Elmer, James Scott
58 ;
59 ; REGISTRATION NUMBER: 36,129
60 ;
61 ; TELECOMMUNICATION INFORMATION:
62 ;
63 ; TELEPHONE: (919) 541-8614
64 ;
65 ; TELEFAX: (919) 541-8689
66 ;
67 ; INFORMATION FOR SEQ ID NO: 4:
68 ;
69 ; SEQUENCE CHARACTERISTICS:
70 ;
71 ; LENGTH: 484 amino acids
72 ;
73 ; TYPE: amino acid

```

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-565-655-4

Query Match 29.0%; Score 42; DB 1; Length 484;  
Best Local Similarity 45.0%; Pred. No. 70;  
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVGAENKDD 26  
||| ||| ||| : : :  
Db 391 LDVSLGSEIKVGVSTYQTD 410

RESULT 11  
US-08-946-967-4  
Sequence 4, Application US/08946967  
Patent No. 5882869

GENERAL INFORMATION:  
APPLICANT: Potter, Sharon L  
TITLE OF INVENTION: Plant Adenylosuccinate Synthetase and  
TITLE OF INVENTION: DNA Coding Therefor  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,967  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-946-967-4

Query Match 29.0%; Score 42; DB 2; Length 484;  
Best Local Similarity 45.0%; Pred. No. 70;  
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVGAENKDD 26  
||| ||| ||| : : :  
Db 391 LDVSLGSEIKVGVSTYQTD 410

RESULT 12  
US-09-231-529-3  
Sequence 3, Application US/09231529  
Patent No. 6096308

GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi  
TITLE OF INVENTION: HUMAN PROTEIN KINASE AND KINASE INHIBITORS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/231,529  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/977,816  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0429 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: KIDNOT25  
CLONE: 353694  
US-09-231-529-3

Query Match 28.6%; Score 41.5; DB 3; Length 365;  
Best Local Similarity 38.5%; Pred. No. 59;  
Matches 10; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

QY 1 KCLG-----YHLDVSLAFSEIS 17  
||:| :||| :|||  
Db 69 KCVGKEDTGSMTMMDVACAFDEIS 94

RESULT 13  
US-08-977-816-3  
Sequence 3, Application US/08977816  
Patent No. 6194186

GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN PROTEIN KINASE AND KINASE INHIBITORS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS

```

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/977,816
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0429 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KIDNOT25
CLONE: 353694
US-08-977-816-3

```

```

Query Match      28.6%; Score 41.5; DB 4; Length 365;
Best Local Similarity 38.5%; Pred. No. 59;
Matches 10; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

```

```

QY 1 KCLG-----YHLDVSLAFSEIS 17
    |||:|||||
Db 69 KCVGKPTGSMATNMMDVACAFDEIS 94

```

```

RESULT 14
US-08-262-220-6
Sequence 6, Application US/08262220
Patent No. 6054296
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,220
FILING DATE: 20-JUN-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BERGSTROM-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 619 amino acids
TYPE: amino acid
TOPOLOGY: linear

```

```

MOLECULE TYPE: protein
US-08-262-220-6

```

```

Query Match      28.6%; Score 41.5; DB 3; Length 619;
Best Local Similarity 45.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 8 DVSLAFSE-LSVGAEPNKDD 26
    ::|||1111111111
Db 372 ELGIATFGASIGLAWNKDD 391

```

```

RESULT 15
US-08-471-733-6
Sequence 6, Application US/08471733
Patent No. 6068842
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,733
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/262,220
FILING DATE: 20-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BERGSTROM-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 619 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-471-733-6

```

```

Query Match      28.6%; Score 41.5; DB 3; Length 619;
Best Local Similarity 45.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;
QY 8 DVSLAFSE-LSVGAEPNKDD 26
    ::|||1111111111
Db 372 ELGIATFGASIGLAWNKDD 391

```

Search completed: January 6, 2002, 09:05:06  
Job time: 194 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:23:06 ; Search time 12.65 seconds  
(without alignments)  
102.369 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFDVTTRENE 17

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 2975

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :  
1: PIR:68:\*  
2: PIR:1:\*  
3: PIR:3:\*  
4: PIR:4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 28    | 29.2        | 15     | 2     | A61612      |
| 2          | 25    | 26.0        | 15     | 2     | PC2215      |
| 3          | 24    | 25.0        | 13     | 2     | H56046      |
| 4          | 23    | 24.0        | 9      | 2     | A11497      |
| 5          | 23    | 24.0        | 14     | 2     | PS0371      |
| 6          | 23    | 24.0        | 17     | 2     | A61117      |
| 7          | 22    | 22.9        | 16     | 2     | A59042      |
| 8          | 22    | 22.9        | 16     | 2     | G49039      |
| 9          | 21    | 21.9        | 12     | 2     | T46794      |
| 10         | 21    | 21.9        | 13     | 2     | S08575      |
| 11         | 21    | 21.9        | 16     | 2     | I52226      |
| 12         | 21    | 20.8        | 8      | 2     | S71919      |
| 13         | 20    | 20.8        | 9      | 2     | A12872      |
| 14         | 20    | 20.8        | 14     | 2     | PT0077      |
| 15         | 20    | 20.8        | 14     | 2     | S29486      |
| 16         | 20    | 20.8        | 15     | 2     | S32677      |
| 17         | 20    | 20.8        | 15     | 2     | PA0009      |
| 18         | 20    | 20.8        | 17     | 2     | D22595      |
| 19         | 19    | 19.8        | 9      | 2     | PT0272      |
| 20         | 19    | 19.8        | 9      | 2     | A42266      |
| 21         | 19    | 19.8        | 10     | 2     | PH0895      |
| 22         | 19    | 19.8        | 11     | 4     | PC2124      |
| 23         | 19    | 19.8        | 14     | 2     | PH1627      |
| 24         | 19    | 19.8        | 15     | 2     | S36889      |
| 25         | 19    | 19.8        | 15     | 2     | PT0095      |
| 26         | 19    | 19.8        | 17     | 2     | S20490      |
| 27         | 18    | 18.8        | 7      | 2     | S08606      |
| 28         | 18    | 18.8        | 10     | 2     | S27873      |
| 29         | 18    | 18.8        | 10     | 2     | PT0284      |

|    |    |      |    |   |        |                     |
|----|----|------|----|---|--------|---------------------|
| 30 | 18 | 18.8 | 12 | 2 | B61497 | seed protein ws-17  |
| 31 | 18 | 18.8 | 13 | 2 | G56046 | urinary tract ston  |
| 32 | 18 | 18.8 | 14 | 2 | A41589 | 25k elastin-bindin  |
| 33 | 18 | 18.8 | 14 | 2 | PA0007 | lectin B1 - Psopho  |
| 34 | 18 | 18.8 | 14 | 2 | 164815 | carbonic anhydrase  |
| 35 | 18 | 18.8 | 14 | 2 | PH1347 | Ig heavy chain DJ   |
| 36 | 18 | 18.8 | 14 | 2 | PH1625 | Ig H chain V-D-J r  |
| 37 | 18 | 18.8 | 15 | 2 | I49407 | placental calcium-  |
| 38 | 18 | 18.8 | 15 | 2 | PA0005 | lectin A1 - Psopho  |
| 39 | 18 | 18.8 | 15 | 2 | PA0006 | lectin A3 - Psopho  |
| 40 | 18 | 18.8 | 15 | 2 | PA0008 | lectin B2 - Psopho  |
| 41 | 18 | 18.8 | 15 | 2 | PT0222 | Ig heavy chain CDR  |
| 42 | 18 | 18.8 | 15 | 2 | PH1310 | Ig heavy chain DJ   |
| 43 | 18 | 18.8 | 16 | 2 | A59046 | alpha-conotoxin MI  |
| 44 | 18 | 18.8 | 16 | 2 | PH1778 | T cell receptor al  |
| 45 | 18 | 18.8 | 17 | 1 | A61339 | vesiculakinin 1 - e |
| 46 | 18 | 18.8 | 17 | 1 | A05168 | conotoxin G [val]   |
| 47 | 18 | 18.8 | 17 | 2 | B61491 | seed protein ws-2   |
| 48 | 18 | 18.8 | 17 | 2 | S69164 | ferredoxin al - Ja  |
| 49 | 17 | 17.7 | 8  | 2 | PC4131 | hypothetical prote  |
| 50 | 17 | 17.7 | 8  | 2 | A25836 | L-serine dehydrata  |

## ALIGNMENTS

RESULT 1  
A61612  
allatostatin - tobacco hornworm  
C:Species: Manduca sexta (tobacco hornworm)  
C:Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 11-Jul-1997  
C:Accession: A61612  
R:Kramer, S.J.; Toschl, A.; Miller, C.A.; Kataoka, H.; Qulstad, G.B.; Li, J.P.; Carne  
Proc. Natl. Acad. Sci. U.S.A. 88, 9458-9462, 1991  
A:Title: Identification of an allatostatin from the tobacco hornworm Manduca sexta.  
A:Reference number: A61612; MUID:92052112  
A:Accession: A61612  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <KRA>  
C:Keywords: neuropeptide; pyrrolidone; carboxylic acid (Gln) #status experimental  
F:1/Modified site: pyrrolidone, carboxylic acid (Gln) #status experimental

Query Match 29.2%; Score 28; DB 2; Length 15;  
Best Local Similarity 40.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 FEHCNFDVTT 10  
| | | | |  
Db 4 FRCQYFNPI 13

RESULT 2  
PC2215  
fibrinogenolytic proteinase A2 (EC 3.4.21.-) - western diamondback rattlesnake (fregm  
N:Alternate names: alpha-fibrinogenase A2  
C:Species: Crotalus atrox (western diamondback rattlesnake)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 17-Mar-1999  
C:Accession: PC2215  
R:Hung, C.C.; Chlou, S.H.  
Biochem. Biophys. Res. Commun. 201, 1414-1423, 1994  
A:Title: Isolation of multiple isoforms of alpha-fibrinogenase from the western diamo  
viper.  
A:Reference number: PC2214; MUID:94296418  
A:Accession: PC2215  
A:Molecule type: protein  
A:Residues: 1-15 <HUN>  
C:Keywords: hydrolase; serine proteinase

Query Match 26.0%; Score 25; DB 2; Length 15;  
Best Local Similarity 42.9%; Pred. No. 6.6e+02;

Matches 3: Conservative 2: Mismatches 2: Indels 0: Gaps 0:  
 QY 2 EHCNND 8  
 : | | :  
 Db 5 DECNINE 11

RESULT 3  
 H56046  
 urinary tract stone matrix protein 10, 42K - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 12-Apr-1995 #sequence\_revision 12-Apr-1995 #text\_change 12-Apr-1995  
 C:Accession: H56046  
 R:Binette, J.P.; Binette, M.B.; Gawinowicz, M.A.; Kendrick, N.  
 submitted to the Protein Sequence Database, February 1995  
 A:Description: Isolation, characterization and sequence of stone proteins.  
 A:Reference number: A56046  
 A:Accession: H56046  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-13 <BIN>

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 2; Length 13;  
 Matches 4: Conservative 1; Mismatches 2; Indels 0; Gaps 0:  
 QY 3 HCNFNDV 9  
 | | | :  
 Db 2 HSYFNDL 8

RESULT 4  
 A11497  
 Transaldolase (EC 2.2.1.2) III - yeast (*Pichia jadinii*) (fragment)  
 C:Species: *Pichia jadinii*; *Candida utilis*  
 C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 30-Sep-1993  
 C:Accession: A11497  
 R:Teolis, O.; Sun, S.C.  
 Arch. Biochem. Biophys. 167, 525-533, 1975  
 A:Title: Isolation of a peptide containing a histidinyI-cysteinyl sequence from the acti  
 A:Reference number: A11497; MUID:75145197  
 A:Accession: A11497  
 A:Molecule type: protein  
 A:Residues: 1-9 <TSO>  
 C:Keywords: transferase

Query Match  
 Best Local Similarity 24.0%; Score 23; DB 2; Length 9;  
 Matches 3: Conservative 0; Mismatches 0; Indels 0; Gaps 0:  
 QY 3 HCN 5  
 | | | :  
 Db 4 HCN 6

RESULT 5  
 PS0371  
 hypothetical protein (psac region) - *Synechococcus* sp. (fragment)  
 C:Species: *Synechococcus* sp.  
 C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
 C:Accession: PS0371  
 R:Rhiel, E.; Stiewelt, V.L.; Gasparich, G.E.; Bryant, D.A.  
 Gene 112, 123-128, 1992  
 A:Title: The psac genes of *Synechococcus* sp. PCC7002 and *Cyanophora paradoxa*: cloning an  
 A:Reference number: JS0694; MUID:92201692  
 A:Accession: PS0371  
 A:Molecule type: DNA  
 A:Residues: 1-14 <RHI>  
 A:Cross-references: GB:M86238; NID:g154574; PIDN:AAA27351.1; PID:g552030

Query Match  
 Best Local Similarity 24.0%; Score 23; DB 2; Length 14;  
 Matches 5: Conservative 1; Mismatches 1; Indels 0; Gaps 0:  
 QY 8 DVTTRLR 14  
 | | | | :  
 Db 7 DVTGRUQ 13

RESULT 6  
 A61117  
 somatostatin precursor processing enzyme (EC 3.4.21.-) - American gooselish (fragment  
 C:Species: *Lophus americanus* (American gooselish)  
 C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 13-Sep-1996  
 C:Accession: A61117  
 R:MacKinn, R.B.; Noe, B.D.; Spiess, J.  
 Endocrinology 129, 2263-2265, 1991  
 A:Title: Identification of a somatostatin-14-generating propeptide converting enzyme  
 A:Reference number: A61117; MUID:92007528  
 A:Accession: A61117  
 A:Molecule type: protein  
 A:Residues: 1-17 <MAC>  
 A:Experimental source: pancreatic islets  
 C:Superfamily: kexin; subtilisin homology  
 C:Keywords: hydrolase; serine protease

Query Match  
 Best Local Similarity 24.0%; Score 23; DB 2; Length 17;  
 Matches 3: Conservative 3; Mismatches 5; Indels 0; Gaps 0:  
 QY 5 FENDVTRLRE 15  
 | | | : :  
 Db 4 NINDIEVMXD 14

RESULT 7  
 A59042  
 alpha-conotoxin Epi - cone shell (*Conus episcopatus*)  
 C:Species: *Conus episcopatus* (bishop's cone)  
 C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 13-Aug-1999  
 C:Accession: A59042  
 R:Loughnan, M.; Bond, T.; Atkins, A.; Cuevas, J.; Adams, D.J.; Broxton, N.M.; Livett,  
 J. Biol. Chem. 273, 15667-15674, 1998  
 A:Title: Alpha-conotoxin Epi, a novel sulfated peptide from *Conus episcopatus* that se  
 A:Reference number: A59042; MUID:98288307  
 A:Accession: A59042  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-16 <LOU>  
 C:Superfamily: alpha-conotoxin  
 C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neu  
 F:1-16/Product: alpha-conotoxin Epi #status experimental <MAT>  
 F:2-8,3-16/Disulfide bonds: #status experimental  
 F:15/Binding site: sulfate (Tyr) (covalent) #status experimental  
 F:16/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match  
 Best Local Similarity 22.9%; Score 22; DB 2; Length 16;  
 Matches 3: Conservative 1; Mismatches 1; Indels 0; Gaps 0:  
 QY 4 CNFND 8  
 | | | :  
 Db 8 CNMNN 12

RESULT 8  
 G49039  
 T-cell receptor beta chain V-D-J-C region (V beta 4, J beta 2.2) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 30-May-1997  
 C:Accession: G49039



Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 HCN 5  
 Db 4 HCB 6

## RESULT 14

PT0077  
 proteochondroitin core protein - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 04-Sep-1998  
 C:Accession: PT0077  
 R:Marcum, J.A.; Thompson, M.A.  
 Biochem. Biophys. Res. Commun. 175, 706-712, 1991  
 A:Title: The amino-terminal region of a proteochondroitin core protein, secreted by aort  
 om human bone.  
 A:Reference number: PT0077; MUID:91207372  
 A:Accession: PT0077  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-14 <MAR>  
 C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan

Query Match 20.8%; Score 20; DB 2; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 4e+03;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FEHCNNDVT 10  
 Db 3 FEOKGFWDFT 12

## RESULT 15

S29486  
 GTP-binding protein o-rab3 - electric ray (Discopyge ommata) (fragment)  
 C:Species: Discopyge ommata  
 C:Date: 22-Nov-1993 #sequence\_revision 27-Feb-1997 #text\_change 13-Mar-1997  
 C:Accession: S29486  
 R:Volkmann, W.; Pevsner, J.; Elferink, L.A.; Scheller, R.H.  
 FEBS Lett. 317, 53-56, 1993  
 A:Title: Association of three small GTP-binding proteins with cholinergic synaptic vesic  
 A:Reference number: S29486; MUID:93154521  
 A:Accession: S29486  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-14 <VOL>

Query Match 20.8%; Score 20; DB 2; Length 14;  
 Best Local Similarity 80.0%; Pred. No. 4e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 NFNDV 9  
 Db 8 NFNAV 12

## RESULT 16

S32677  
 nitrogenase cofactor synthesis protein nifs - Anabaena variabilis (fragment)  
 N:Contains: L-cysteine sulfurtransferase (EC 2.8.1.-)  
 C:Species: Anabaena variabilis  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 03-Dec-1999  
 C:Accession: S32677  
 R:Komerjahn, U.; Boelme, H.  
 submitted to the EMBL Data Library, December 1992  
 A:Description: Cloning and expression in E. coli of the Anabaena.  
 A:Reference number: S32675  
 A:Accession: S32677  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-15 <MON>  
 A:Cross-references: EMBL:X69898; NID:q296503; PIDN:CAA49523.1; PID:q296506  
 C:Superfamily: nitrogen fixation protein nifs  
 C:Keywords: sulfurtransferase

Query Match 20.8%; Score 20; DB 2; Length 15;  
 Best Local Similarity 42.9%; Pred. No. 4.3e+03;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
 Db 8 NNATKTV 14

## RESULT 17

PA0009  
 seed storage protein beta-chain 7 - Arabidopsis thaliana (fragment)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 23-Mar-1995  
 C:Accession: PA0009  
 R:Kano, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
 submitted to JIPID, July 1994  
 A:Description: Separation and characterization of Arabidopsis proteins by two-dimensi  
 A:Reference number: PA0001  
 A:Accession: PA0009  
 A:Molecule type: protein  
 A:Residues: 1-15 <KAM>  
 A:Experimental source: seed  
 C:Keywords: seed; storage protein

Query Match 20.8%; Score 20; DB 2; Length 15;  
 Best Local Similarity 42.9%; Pred. No. 4.3e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 TTRAREN 16  
 Db 8 TAKIHEN 14

## RESULT 18

D22595  
 bombolitin IV - American common bumblebee  
 C:Species: Bombus pennsylvanicus (American common bumblebee)  
 C:Date: 29-Aug-1987 #sequence\_revision 29-Aug-1987 #text\_change 08-Dec-1995  
 C:Accession: D22595  
 R:Argiolas, A.; Pisano, J.J.  
 J. Biol. Chem. 260, 1437-1444, 1985  
 A:Title: Bombolitins, a new class of mast cell degranulating peptides from the venom  
 A:Reference number: A92504; MUID:85105003  
 A:Accession: D22595  
 A:Molecule type: protein  
 A:Residues: 1-17 <ARG>  
 C:Keywords: amidated carboxyl end; hemolysis; venom  
 F:17/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 20.8%; Score 20; DB 2; Length 17;  
 Best Local Similarity 33.3%; Pred. No. 4.8e+03;  
 Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 NFNDVTTRL 13  
 Db 2 NIKDILAKL 10

## RESULT 19

PT0272  
 Ig heavy chain CRD3 region (clone 3-103B) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0272



R.Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: P10222; MUID:91108337  
 A:Accession: P10272  
 A:Molecule type: DNA  
 A:Residues: 1-9 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotrimer; immunoglobulin

Query Match 19.8%; Score 19; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 NFND 8  
 DB 2 NMND 5

RESULT 20  
 A42266

peptidylglycine monooxygenase (EC 1.14.17.3), rPAM-5 - rat (fragment)  
 N:Alternate names: peptidylglycine alpha-amidating monooxygenase  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-May-1995  
 C:Accession: A42266  
 R:Elipper, B.A.; Green, C.B.; Campbell, T.A.; Stoffers, D.A.; Keutmann, H.T.; Mains, R.E.  
 J. Biol. Chem. 267, 4008-4015, 1992  
 A:Title: Alternative splicing and endoproteolytic processing generate tissue-specific fo  
 A:Reference number: A42266; MUID:92156145  
 A:Accession: A42266  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-9 <EIP>  
 A:Experimental source: pituitary  
 A:Note: sequence extracted from NCBI backbone (NCBIN:82733, NCBIPI:82750)  
 C:Keywords: oxidoreductase

Query Match 19.8%; Score 19; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+05;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 HCNFND 8  
 DB 2 HICFND 7

RESULT 21  
 PH0895

T-cell receptor beta chain V-D-J region - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
 C:Accession: PH0895; PH0896  
 R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
 J. Exp. Med. 174, 1467-1476, 1991  
 A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy  
 A:Reference number: PH0891; MUID:92078857  
 A:Accession: PH0895  
 A:Molecule type: mRNA  
 A:Residues: 1-10 <GOL>  
 A:Experimental source: myelin basic protein-immunized T-cell; clones 15 and 16  
 C:Keywords: T-cell receptor

Query Match 19.8%; Score 19; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 4.1e+03;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 CNFNDYTRL 13  
 DB 1 CASDSSERL 10

RESULT 22  
 PC2124  
 aminotransferase chimera DY376 - synthetic (fragment)

C:Species: Synthetic  
 C:Date: 28-May-1999 #sequence\_revision 28-May-1999 #text\_change 28-May-1999  
 C:Accession: PC2124  
 R: Miyazawa, K.; Kawaguchi, S.; Okamoto, A.; Kato, R.; Ogawa, T.; Kuramitsu, S.  
 J. Biochem. 115, 568-577, 1994  
 A:Title: Construction of aminotransferase chimeras and analysis of their substrate sp

A:Reference number: JX0315; MUID:94334304

A:Accession: PC2124  
 A:Molecule type: DNA  
 A:Residues: 1-11 <MTY>  
 C:Comment: This is a chimeric enzyme of Escherichia coli aspartate aminotransferase (C:gene: aspC; tyrb  
 C:Comment: The parental enzymes catalyze the reversible amino group transfer reaction  
 C:Keywords: aminotransferase

Query Match 19.8%; Score 19; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 RLRE 15  
 DB 2 RLRE 5

RESULT 23  
 PH1627

Ig H chain V-D-J region (clone B-less 120) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
 C:Accession: PH1627  
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.  
 J. Exp. Med. 178, 317-329, 1993  
 A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m  
 A:Reference number: PH1580; MUID:93301609  
 A:Accession: PH1627  
 A:Molecule type: DNA  
 A:Residues: 1-14 <LEV>  
 A:Experimental source: bone marrow pre-B lymphocyte  
 C:Keywords: immunoglobulin

Query Match 19.8%; Score 19; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 VTTR 12  
 DB 5 VTTR 8

RESULT 24  
 S36889

Ribosomal protein S20 - Mycobacterium bovis (fragment)  
 C:Species: Mycobacterium bovis  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
 C:Accession: S36889  
 R:Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.  
 FEBS Lett. 331, 9-14, 1993  
 A:Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Myco

A:Reference number: S36887; MUID:94009653  
 A:Accession: S36889  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <OHA>

C:Keywords: protein biosynthesis; ribosome

Query Match 19.8%; Score 19; DB 2; Length 15;  
 Best Local Similarity 38.5%; Pred. No. 6.2e+03;  
 Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 5 NENDVTRLRENE 17  
 | | | | |  
 Db 2 NIKSOKRNRRTNE 14

## RESULT 25

PT0095

H+-transporting ATP synthase (EC 3.6.1.34) gamma chain, mitochondrial - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 21-Aug-1998

C:Accession: PT0095

R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.

A:Description: Proteome analysis of mouse brain.

A:Reference number: PT0091

A:Accession: PT0095

A:Molecule type: protein

A:Residues: 1-15 &lt;KAW&gt;

A:Experimental source: brain, striatum

C:Keywords: hydrolase; mitochondrion

## Query Match

Best Local Similarity 19.8%; Score 19; DB 2; Length 15;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 VTTRLR 14

: | | | |

Db 6 ITXRLK 11

Search completed: April 2, 2002, 09:25:00  
 Job time: 114 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:24:41 : Search time 10.02 seconds  
(without alignments)  
62.206 Million cell updates/sec

Title: us-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVYTRLENE 17

Scoring table:  
BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 836

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 28    | 29.2        | 15     | 1     | ALIS_MANSE  |
| 2          | 23    | 24.0        | 9      | 1     | TAL3_PICJA  |
| 3          | 23    | 24.0        | 17     | 1     | JHRP_PLAVG  |
| 4          | 22    | 22.9        | 10     | 1     | UPA9_HUMAN  |
| 5          | 22    | 22.9        | 16     | 1     | CMAL_CONEP  |
| 6          | 21    | 21.9        | 15     | 1     | MCA2_RHOOP  |
| 7          | 21    | 21.9        | 15     | 1     | UC01_MAIZE  |
| 8          | 21    | 21.9        | 15     | 1     | UC13_MAIZE  |
| 9          | 20    | 20.8        | 9      | 1     | OXYF_SCYCA  |
| 10         | 20    | 20.8        | 9      | 1     | TAL1_PICJA  |
| 11         | 20    | 20.8        | 10     | 1     | URA7_HUMAN  |
| 12         | 20    | 20.8        | 15     | 1     | NIS1_ANASQ  |
| 13         | 20    | 20.8        | 17     | 1     | BOL4_MEGPE  |
| 14         | 19    | 19.8        | 12     | 1     | URA_SCYCA   |
| 15         | 19    | 19.8        | 14     | 1     | UN46_CLOPA  |
| 16         | 19    | 19.8        | 15     | 1     | ITRB_ALBUJ  |
| 17         | 19    | 19.8        | 16     | 1     | FIBA_MUSVI  |
| 18         | 18    | 18.8        | 8      | 1     | ALU1_CYPDO  |
| 19         | 18    | 18.8        | 8      | 1     | LMT2_LOCM1  |
| 20         | 18    | 18.8        | 12     | 1     | HCTB_MEGCR  |
| 21         | 18    | 18.8        | 13     | 1     | ORCK_ORCLT  |
| 22         | 18    | 18.8        | 13     | 1     | UP51_UPEIN  |
| 23         | 18    | 18.8        | 14     | 1     | LECB_PSOSC  |
| 24         | 18    | 18.8        | 15     | 1     | LECI_PSOSC  |
| 25         | 18    | 18.8        | 15     | 1     | LEC2_PSOSC  |
| 26         | 18    | 18.8        | 15     | 1     | LEC3_PSOSC  |
| 27         | 18    | 18.8        | 16     | 1     | CXA2_CONMA  |
| 28         | 18    | 18.8        | 17     | 1     | VESP_VESWC  |
| 29         | 17    | 17.7        | 9      | 1     | UPA6_HUMAN  |
| 30         | 17    | 17.7        | 13     | 1     | LMT4_LOCM1  |
| 31         | 17    | 17.7        | 15     | 1     | CXAL_CONGE  |
| 32         | 17    | 17.7        | 16     | 1     | AHL_PRUSE   |
| 33         | 17    | 17.7        | 16     | 1     | DBH3_RHILE  |

## ALIGNMENTS

| RESULT                | ID  | ALIS_MANSE | STANDARD | PRT | 15 AA |
|-----------------------|---|------------|----------|-----|-------|
| AC                    | P42559  |            |          |     |       |
| DT                    | 01-NOV-1995 (Rel. 32, Created)  |            |          |     |       |
| DT                    | 01-NOV-1995 (Rel. 32, Last sequence update)   |            |          |     |       |
| DT                    | 30-MAY-2000 (Rel. 39, Last annotation update)   |            |          |     |       |
| DE                    | ALATOSTRATIN (MAS-AS).  |            |          |     |       |
| OS                    | Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  |            |          |     |       |
| OC                    | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;   |            |          |     |       |
| OC                    | Phryganea; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  |            |          |     |       |
| OC                    | Sphingidae; Sphingidae; Sphinginae; Manduca.  |            |          |     |       |
| OX                    | NCBI_TaxID=7130;  |            |          |     |       |
| RN                    | [1]   |            |          |     |       |
| RP                    | SEQUENCE:   |            |          |     |       |
| RC                    | TISUE=Head;   |            |          |     |       |
| RX                    | MEDLINE=92052112; PubMed=1946359;   |            |          |     |       |
| RA                    | Kramer S.J., Toschl A., Miller C.A., Kataoka H., Quistad G.B.,  |            |          |     |       |
| RA                    | Li J.P., Carney R.L., Schooley D.A.;  |            |          |     |       |
| RT                    | "Identification of an allatostatin from the tobacco hornworm Manduca sexta."  |            |          |     |       |
| RL                    | Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).   |            |          |     |       |
| CC                    | -1- FUNCTION: STRONGLY INHIBITS JUVENILE HORMONE BIOSYNTHESIS IN VITRO BY THE CORPORA ALLATA FROM FIFTH-STADIUM LARVAE AND ADULT FEMALES. |            |          |     |       |
| CC                    | -1- SIMILARITY: BELONGS TO THE ALATOSTRATIN FAMILY.   |            |          |     |       |
| KW                    | Neuropeptide. 1   |            |          |     |       |
| FT                    | MOD_RES 1   |            |          |     |       |
| FT                    | MOD_RES 15 AA; 1908 MM; PYROLIDONE CARBOXYLIC ACID.   |            |          |     |       |
| SQ                    | SEQUENCE 15 AA; 1605877CDBEC838E CRC64;   |            |          |     |       |
| Query Match           | 29.2%; Score 28; DB 1; Length 15;   |            |          |     |       |
| Best Local Similarity | 40.0%; Pred. No. 75;  |            |          |     |       |
| Matches               | 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  |            |          |     |       |
| QY                    | 1 FEHCNFNDVY 10   |            |          |     |       |
| DB                    | 4 FROCYFNPIS 13   |            |          |     |       |
| RESULT                | 2   |            |          |     |       |
| ID                    | TAL3_PICJA  |            |          |     |       |
| AC                    | P17441  |            |          |     |       |
| DT                    | 01-AUG-1990 (Rel. 15, Created)  |            |          |     |       |
| DT                    | 01-FEB-1994 (Rel. 28, Last sequence update)   |            |          |     |       |
| DE                    | TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).  |            |          |     |       |
| OC                    | Pichia jadinii (Yeast) (Candida utilis).  |            |          |     |       |
| OC                    | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  |            |          |     |       |
| OC                    | Saccharomycetales; Saccharomycetaceae; Pichia.  |            |          |     |       |
| OX                    | NCBI_TaxID=4903;  |            |          |     |       |

RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=75145197; PubMed=1092268;  
 RA Tsolas O., Sun S.C.;  
 RT "Isolation of a peptide containing a histidyl-cysteiny sequence  
 from the active center of transaldolase.";  
 RL Arch. Biochem. Biophys. 167:525-533(1975).  
 CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
 METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
 CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE  
 3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.  
 DR PIR: A11497; A11497.  
 DR InterPro: IPR001585; Transaldolase.  
 DR PROSITE: PS00958; TRANSALDOLASE\_2; PARTIAL.  
 DR PROSITE: PS01054; TRANSALDOLASE\_1; PARTIAL.  
 KW transferase; Pentose shunt.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 9 AA; 1033 MW; 325A31A44EB1E058 CRC64;

Query Match 24.0%; Score 23; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HCN 5  
 Db 4 HCN 6

RESULT 3  
 ID JHBP\_PLAVG STANDARD; PRT; 17 AA.  
 AC P56675;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 GN JUVENILE HORMONE-BINDING PROTEIN (FRAGMENT).  
 OS Platypreria virginialis (Ranchman's tiger moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Noctuidae; Actiidae; Platypreria.  
 OX NCBI\_TaxID=30227;  
 RN [1]  
 RP SEQUENCE.  
 RA Prestwich G.D., Atkinson J.K.;  
 RT "Rapid purification and N-terminal amino acid sequence of a  
 photoaffinity-labeled juvenile hormone binding protein from an arctiid  
 moth larva, Platypreria virginialis.";  
 RL Insect Biochem. 20:801-807(1990).  
 CC -1- FUNCTION: PREVENTS JUVENILE HORMONE FROM BEING HYDROLYZED BY  
 CC GENERAL ESTERASES BY COMBINING WITH IT SPECIFICALLY.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 FT NON\_TER 17  
 FT NON\_TER 17  
 SO SEQUENCE 17 AA; 1880 MW; B1205F0977320A8A CRC64;

Query Match 24.0%; Score 23; DB 1; Length 17;  
 Best Local Similarity 33.3%; Pred. No. 5.9e+02;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 FERCHENDV 9  
 Db 4 FDPGSTODI 12

RESULT 4  
 ID UP9A\_HUMAN STANDARD; PRT; 10 AA.  
 AC P30095;

DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 35) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097.  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquall C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RL "Plasma protein map: an update by microsequencing."  
 CC Electrophoresis 13:707-714(1992).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 7.2. ITS MW IS: 15 KDA.  
 DR SWISS-2DPAGE: P30095; HUMAN.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 10 AA; 1233 MW; 37AD72B409C681B7 CRC64;

Query Match 22.9%; Score 22; DB 1; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 4.9e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 TLRENE 17  
 Db 4 TMFRNE 10

RESULT 5  
 ID CXAL\_CONEP STANDARD; PRT; 16 AA.  
 AC P56638;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 GN ALPHA-CONOTOXIN EPI.  
 OS Conus episcopatus (Bishop's cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conidae; Conus.  
 OX NCBI\_TaxID=88764;  
 RN [1]  
 RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).  
 RX MEDLINE=96376423; PubMed=9708977.  
 RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,  
 RA Alewood P.F., Lewis R.J., Martin J.L.;  
 RT "The 1.1-A resolution crystal structure of [tyr15]EPI, a novel  
 alpha-conotoxin from Conus episcopatus, solved by direct methods.";  
 RL Biochemistry 37:11425-11433(1998).  
 CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACR) AND THUS  
 CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC  
 CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-3/  
 CC 3/BETA-4 SUBUNITS.  
 DR PDB: 1A0M; 13-JAN-99.  
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
 KW Venom; Sulfation; 3D-structure.  
 FT DISULFID 2  
 FT DISULFID 3  
 FT DISULFID 8  
 FT MOD\_RES 15  
 FT MOD\_RES 15  
 FT MOD\_RES 16  
 FT MOD\_RES 16  
 SO SEQUENCE 16 AA; 1792 MW; C6385F376C9B4C CRC64;

Query Match 22.9%; Score 22; DB 1; Length 16;  
 Best Local Similarity 60.0%; Pred. No. 8.1e+02;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNFND 8  
1 1 1  
DB 8 CNMNN 12

## RESULT 6

MCA2\_RHOOP STANDARD; PRT; 15 AA.  
ID MCA2\_RHOOP  
AC P56870;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PUTATIVE MALEYLACETATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).  
OS Rhodococcus opacus (Nocardia opaca).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriineae; Nocardiaceae; Rhodococcus.  
OX NCBI\_TaxID=37919;  
RN [1]  
RP SEQUENCE.

RA STRAIN-1CP;  
RC MEDLINE-98324954; PubMed-9657989;  
RA Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;  
RT "Characterization of a maleylacetate reductase encoding region from  
Rhodococcus opacus 1CP.";  
RL J. Bacteriol. 180:3503-3508(1998).  
CC -1- CATALYTIC ACTIVITY: 3-OXOADIPATE + NAD(P)(+) = 2-MALEYLACETATE +  
NAD(P)H.  
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).  
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC  
AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL  
PRODUCTS AND AS INDUSTRIAL EFFLUENT.  
CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE  
FAMILY.  
DR InterPro: IPR001670; Fe-ADH.  
DR PROSITE: PS00913; ADH\_IRON\_1; PARTIAL.  
DR PROSITE: PS00060; ADH\_IRON\_2; PARTIAL.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA: 1884 MW: 58DA90DD038F025E CRC64;

Query Match 21.9%; Score 21; DB 1; Length 15;  
Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FEHCN 5  
1 1 1  
DB 3 FEHEN 7

## RESULT 7

UC01\_MAIZE STANDARD; PRT; 15 AA.  
ID UC01\_MAIZE  
AC P80607;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 113)  
(FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.

RA Tissue-Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
RA Perrolliet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
genome analysis program.";  
RT Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN

CC PROTEIN IS: 6.3, ITS MW IS: 37.4 KDA.  
DR Maize-2DPAGE; P80607; COLEOPTILE.  
DR MaizeDB: 123920; -;  
FT NON\_TER 1 1  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA: 1844 MW: 6F73A6DDCA0F05A5 CRC64;

Query Match 21.9%; Score 21; DB 1; Length 15;  
Best Local Similarity 44.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 FEHCN 5  
1 1 1  
DB 7 YELYNRNDI 15

## RESULT 8

UC13\_MAIZE STANDARD; PRT; 15 AA.  
ID UC13\_MAIZE  
AC P80619;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 243)  
(FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.

RA Tissue-Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
RA Perrolliet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
genome analysis program.";  
RT Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
PROTEIN IS: 6.8, ITS MW IS: 56.9 KDA.  
CC -1- SIMILARITY: TO GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASES.  
DR Maize-2DPAGE; P80619; COLEOPTILE.  
DR MaizeDB: 123946; -;  
FT NON\_TER 1 1  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA: 1739 MW: 02038BE7471AE038 CRC64;

Query Match 21.9%; Score 21; DB 1; Length 15;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 7 NDVTRLRNE 17  
1 1 1  
DB 4 NEVVIRLOXSE 14

## RESULT 9

OXYF\_SCYCA STANDARD; PRT; 9 AA.  
ID OXYF\_SCYCA  
AC P42997;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE PHASVATOCIN.  
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;  
OC Scyllorhinidae; Scyllorhinus.  
OX NCBI\_TaxID=7830;  
RN [1]  
RP SEQUENCE.

RC TISSUE=Pituitary;  
 RX MEDLINE=95062247; PubMed-7972045;  
 RA Chauvet J., Rouille Y., Chauveau C., Chauvet M.-T., Acher R.;  
 RT "Special evolution of neurohypophyseal hormones in cartilaginous  
 RT fishes: aspartic acid and phasvatoxin, two oxytocin-like peptides  
 RT isolated from the spotted dogfish, *Scyliorhinus caniculus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).  
 CC -1- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERUS.  
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR InterPro: IPR000981; Neurohypophys\_horm.  
 DR Pfam: PF00220; hormone; 1  
 DR PROSITE: PS00264; Neurohypophys\_HORM; 1.  
 KW Hormone: Amidation.  
 FT Disulfid  
 FT MOD\_RES 9 6  
 FT SEQUENCE 9 AA; 1016 MW; 17EDD76EB4449DB CRC64;

Query Match 20.8%; Score 20; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 1e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CNFND 8  
 DB 1 CYFNN 5

RESULT 10  
 ID TALL\_PICJA STANDARD; PRT; 9 AA.  
 AC P17440;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE TRANSALDOLASE I (EC 2.2.1.2) (FRAGMENT).  
 OS Pichia jadinii (Yeast) (Candida utilis).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Pichia.  
 NX NCBI\_TaxID=4903;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77110646; PubMed=556924;  
 RA Sun S.C., Joris L., Tsolas O.;  
 RT "Purification of crystallization of transaldolase isozyme I and  
 RT evidence for different genetic origin of isozymes I and III in  
 RT Candida utilis.";  
 RL Arch. Biochem. Biophys. 178:69-78(1977).  
 CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
 CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
 CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE  
 CC 3-PHOSPHATE -> D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.  
 CC -1- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.  
 DR PIR: A12872; A12872.  
 DR InterPro: IPR001585; Transaldolase.  
 DR PROSITE: PS00958; TRANSALDOLASE\_2; PARTIAL.  
 DR PROSITE: PS01054; TRANSALDOLASE\_1; PARTIAL.  
 KW Transferase; Pentose shunt.  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 FT SEQUENCE 9 AA; 1008 MW; 274EF31AF0EB1E058 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 HCN 5  
 DB 4 HCB 6

RESULT 11

URAT\_HUMAN  
 ID URAT\_HUMAN STANDARD; PRT; 10 AA.  
 AC P34990;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF RED BLOOD CELLS (SPOT 2D-006H0)  
 OS (FRAGMENT).  
 DE Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Erythrocyte;  
 RA Golaz O., Hughes G.J., Frutiger S., Paquet N., Bairoch A.,  
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Appel R.D., Walzer C.,  
 RA Balant L., Hochstrasser D.F.;  
 RL Submitted (FEB-1994) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.77, ITS MW IS: 26 KDA.  
 DR SWISS-2DPAGE; P34990; HDMAN.  
 FT NON\_TER 10  
 FT SEQUENCE 10 AA; 1161 MW; F72D82E9C44B1871 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NFNDV 9  
 DB 6 NFNDI 10

RESULT 12  
 ID NISL\_ANASO STANDARD; PRT; 15 AA.  
 AC 044507;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CYSTEINE DESULFURASE 1 (EC 4.4.1.-) (NITROGENASE METALLOCLUSTERS  
 DE BIOSYNTHESIS PROTEIN NIFS1) (FRAGMENT).  
 GN NIFS1 OR NIFS.  
 OS Anabaena sp. (strain PCC 7937 / ATCC 29413).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 NX NCBI\_TaxID=1172;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Monnerjahn U., Boelme H.;  
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP EXPRESSION PATTERN.  
 RX MEDLINE=96016168; PubMed=7568132;  
 RA Thiel T., Lyons E.M., Erker J.C., Ernst A.;  
 RT "A second nitrogenase in vegetative cells of a heterocyst-forming  
 RT cyanobacterium";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362(1995).  
 CC -1- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR ATOMS FROM  
 CC CYSTEINE TO PRODUCE ALANINE. SEEMS TO PARTICIPATE IN THE  
 CC BIOSYNTHESIS OF THE NITROGENASE METALLOCLUSTERS BY PROVIDING THE  
 CC INORGANIC SULFUR REQUIRED FOR THE FE-S CORE FORMATION (BY  
 CC SIMILARITY).  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: BELONGS TO THE NIFI GENE CLUSTER WHICH IS EXPRESSED  
 CC IN HETEROCYST UNDER ANAEROBIC AND AEROBIC CONDITIONS.  
 CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X69898; CAA9523.1; -  
DR InterPro; IPR000192; AminoTransf\_class-V.  
DR PROSITE; PS00595; AA\_TRANSF\_CLASS\_5; PARTIAL.  
KM Nitrogen fixation; Lyase; Pyridoxal phosphate.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1684 MW; 08B8F106DE65547D CRC64;

Query Match 20.8%; Score 20; DB 1; Length 15;  
Best Local Similarity 42.9%; Pred. No. 1.9e+03;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
ID BOLA\_MEGPE STANDARD; PRT; 17 AA.  
Db 8 NNATTKV 14

RESULT 13  
BOLA\_MEGPE  
ID BOLA\_MEGPE STANDARD; PRT; 17 AA.  
AC P07495;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DE BOMBOLITIN IV.  
OS Megabombus pennsylvanicus (American common bumblebee).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;  
OC Apoidea; Apidae; Bombus.  
OX NCBI\_TaxId=28643;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=85105003; PubMed=2578459;  
RA Argiolas A., Pisano J.J.;  
RT Bombolitin, a new class of mast cell degranulating peptides from  
RT the venom of the bumblebee Megabombus pennsylvanicus.";  
RL J. Biol. Chem. 260:1437-1444(1985).  
CC -1- FUNCTION: MAST CELL DEGRANULATING PEPTIDE.  
DR PTR: D22595; D22595.  
KM Mast cell degranulation; Venom.  
SQ SEQUENCE 17 AA; 1873 MW; A34A43514BCFDFB6 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 17;  
Best Local Similarity 33.3%; Pred. No. 1.9e+03;  
Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 NFNDVTTRL 13  
ID UR2\_SCYCA STANDARD; PRT; 12 AA.  
Db 2 NIKDILAKL 10

RESULT 14  
UR2\_SCYCA  
ID UR2\_SCYCA STANDARD; PRT; 12 AA.  
AC P35480;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE UROTENSIN II (U-II) (U11).  
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeidae; Carcharhiniformes;  
OC Scyllorhinidae; Scyllorhinus.  
OX NCBI\_TaxId=7830;  
RN [1]

RP SEQUENCE.  
RC TISSUE=Spinal cord;  
RX MEDLINE=92319231; PubMed=1620290;  
RA Conlon J.M., O'Harte F., Smith D.D., Balmert R.J., Hazon N.;  
RT "Purification and characterization of urotensin II and parvalbumin  
RT from an elasmobranch fish, Scyllorhinus canicula (common dogfish).";  
RL Neuroendocrinology 55:230-235(1992).

CC -1- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A  
CC CORTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH  
CC MUSCLE STIMULATION.  
CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.

DR InterPro; IPR001483; Urotensin\_II.  
DR Pfam; PF02083; Urotensin\_II; 1.  
DR PROSITE; PS00984; UROTENSIN\_II; 1.  
KM Hormone.  
FT DISULFID 6 11  
SQ SEQUENCE 12 AA; 1526 MW; 804729F9D579CEBA CRC64;

Query Match 19.8%; Score 19; DB 1; Length 12;  
Best Local Similarity 75.0%; Pred. No. 1.9e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 NFND 8  
ID UN46\_CLOPA STANDARD; PRT; 14 AA.  
Db 2 NFSD 5

RESULT 15  
UN46\_CLOPA  
ID UN46\_CLOPA STANDARD; PRT; 14 AA.  
AC P81362;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE UNKNOWN PROTEIN CP 46 FROM 2D-PAGE (FRAGMENT).  
OS Clostridium pasteurianum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxId=1501;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=W5;  
RX MEDLINE=98291870; PubMed=9629918;  
RA Flengrind R., Skjeldal L.;  
RT "Two-dimensional gel electrophoresis separation and N-terminal  
RT sequence analysis of proteins from Clostridium pasteurianum W5.";  
RL Electrophoresis 19:802-806(1998).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 5.4, ITS MW IS: 38.2 KDA.  
FT NON\_TER 14 14  
SQ SEQUENCE 14 AA; 1550 MW; 198078F4C0367170 CRC64;

Query Match 19.8%; Score 19; DB 1; Length 14;  
Best Local Similarity 75.0%; Pred. No. 2.2e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 FNDV 9  
ID ITRB\_ALBU STANDARD; PRT; 15 AA.  
Db 3 FNDL 6

RESULT 16  
ITRB\_ALBU  
ID ITRB\_ALBU STANDARD; PRT; 15 AA.  
AC P24927;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE TRYPSIN INHIBITOR B CHAIN (FRAGMENT).  
OS Albizzia julibrissin (Silk tree).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucotsids I; Fabales; Fabaceae; Mimosoidae; Ingeae; Albizia;  
 RN NCBL\_TaxID=3813;  
 [1]  
 RP SEQUENCE.  
 RC MEDLINE=80115605; PubMed=528539;  
 RX MEDLINE=80115605; PubMed=528539;  
 RA Odani S., Ono T., Ikenaka T.;  
 RT "Proteinase inhibitors from a mimosoidae legume, Albizzia  
 julibrissin. Homologues of soybean trypsin inhibitor (Kunitz).";  
 RL J. Biochem. 86:1795-1805(1979).  
 CC -1- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.  
 CC -1- SUBUNIT: HETERODIMER OF AN "A" AND A "B" CHAIN LINKED BY A  
 CC DISULFIDE BOND.  
 CC -1- SIMILARITY: TO BETA CHAINS OF SOME OTHER LEGUMINOUS KUNITZ-TYPE  
 CC INHIBITORS.  
 KW Serine protease inhibitor.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1705 MW; 53165F7E9C45B4D0 CRC64;

Query Match  
 Best Local Similarity 19.8%; Score 19; DB 1; Length 15;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 EHC 4  
 DB 3 DHC 5

RESULT 17  
 FIBA\_MUSVI  
 ID FIBA\_MUSVI STANDARD; PRT; 16 AA.  
 AC P14458;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-JAN-1990 (Rel. 13, Last annotation update)  
 DE FIBRINOPEPTIDE A.  
 OS Mustela vison (American mink).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustela.  
 OX NCBL\_TaxID=9667;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals.";  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION. YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 CC Blood coagulation; Plasma.  
 KW NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1679 MW; 09432A57919ECF66 CRC64;

Query Match  
 Best Local Similarity 19.8%; Score 19; DB 1; Length 16;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 11 TRLENE 17  
 DB 1 TWNESE 7

RESULT 18  
 ALL1\_CYPDPO  
 ID ALL1\_CYPDPO STANDARD; PRT; 8 AA.

AC P82152;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYDIASTATTIN 1.  
 OS Cydia pomonella (Coddling moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 OX NCBL\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Davey H., Johnson A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
 RT "Lepidopteran peptides of the allatostatin superfamily.";  
 RL Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATTIN FAMILY.  
 CC Neuropeptide; Amidation.  
 KW MOD\_RES 8 8  
 FT SEQUENCE 8 AA; 934 MW; C82879C45B51F775 CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 8;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 HCNF 6  
 DB 3 HYNF 6

RESULT 19  
 LMT2\_LOCM1  
 ID LMT2\_LOCM1 STANDARD; PRT; 8 AA.  
 AC P22396;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-AUG-1991 (Rel. 19, Last annotation update)  
 DE LOCUSTAMYTROPIN 2 (LOM-MT-2).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;  
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.  
 OX NCBL\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Corpora cardiaca;  
 RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;  
 RT "Isolation, identification and synthesis of locustamytotropin II, an  
 RT additional neuropeptide of Locusta migratoria. Member of the  
 RT cephalomyotropic peptide family.";  
 RL Insect Biochem. 20:479-484(1990).  
 CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY  
 CC (MYOTROPIC ACTIVITY).  
 CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 CC InterPro: IPR001484; Pyrokinin.  
 DR PROSITE: PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin.  
 FT MOD\_RES 8 8  
 SQ SEQUENCE 8 AA; 934 MW; 2634171A9CAAB7B CRC64;

Query Match  
 Best Local Similarity 18.8%; Score 18; DB 1; Length 8;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 8 DVTRL 13  
 DB 3 DTPRL 8



```

RESULT 20
HCYB_MEGCR STANDARD; PRT; 12 AA.
AC 010584:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEMOCYANIN B CHAIN (KLH-B) (FRAGMENT).
OS Megathura crenulata (Giant keyhole limpet).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC Patellogastropoda; Megathura.
OC NCB1_TaxID=53429;
RN [1]
RP SEQUENCE.
RX MEDLINE=96208935; PubMed=8829804;
RA Swerdlow R.D., Ebert R.F., Lee P., Bonaventura C., Miller K.I.;
RT "Keyhole limpet hemocyanin: structural and functional
characterization of two different subunits and multimers.";
RL Comp. Biochem. Physiol. 113B:537-548(1996).
CC -1- FUNCTION: HEMOCYANINS ARE COPPER-CONTAINING OXYGEN CARRIERS
OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND
ARTHRPODS.
CC -1- SUBUNIT: DIDEAMERS AND EXTENDED MULTIMERS.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: HEMOLYMPH.
CC -1- BIOTECHNOLOGY: POTENT IMMUNOGEN USED CLASSICALLY AS A CARRIER
PROTEIN FOR HAPTENS AND MORE RECENTLY IN HUMAN VACCINES AND FOR
IMMUNOTHERAPY OF BLADDER CANCER.
CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HEMOCYANIN
CC SUBFAMILY.
DR InterPro: IPR000896; Hemocyanin.
DR PROSITE: PS00209; HEMOCYANIN_1; PARTIAL.
DR PROSITE: PS00210; HEMOCYANIN_2; PARTIAL.
KW Respiratory protein; Oxygen transport; Copper; Glycoprotein;
KW Hemolymph.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1345 MW; CBFEBAA44A432412 CRC64;

Query Match 18.8%; Score 18; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLEN 16
Db 1 VDTVRKN 8

RESULT 21
ORCK_ORCLI STANDARD; PRT; 13 AA.
AC P37086:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ORCOKININ.
OS Orconectes limosus (Spinycheek crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Astacidea; Cambaridae; Orconectes.
OC NCB1_TaxID=28379;
RN [1]
RP SEQUENCE.
RX MEDLINE=93126144; PubMed=1480511;
RA Strangler J., Hilbich C., Burdzik S., Keller R.;
RT Orckokin: a novel myotropic peptide from the nervous system of the
crayfish, Orconectes limosus.";
RL Peptides 13:859-864(1992).
CC -1- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS.
CC -1- TISSUE SPECIFICITY: ABDOMINAL NERVE CORD AND HINDGUT.
KW Neuropeptide.

```

```

SQ SEQUENCE 13 AA; 1518 MW; 8A31807BA93A40A CRC64;

Query Match 18.8%; Score 18; DB 1; Length 13;
Best Local Similarity 40.0%; Pred. No. 3e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 NFNDV 9
Db 1 NFDEI 5

RESULT 22
UP51_UPEIN STANDARD; PRT; 13 AA.
ID UP51_UPEIN
AC P82036:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE UPERIN 5.1.
OS Uperoleia inundata (Floodplain toadlet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Myobatrachidae;
OC Uperoleia.
OC NCB1_TaxID=104953;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=SKIN secretion;
RA Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA Adams G.W., Severini C.;
RT "Novel uperin peptides from the dorsal glands of the Australian
RT floodplain toadlet Uperoleia inundata.";
RL Aust. J. Chem. 49:475-484(1996).
CC -1- FUNCTION: UNKNOWN.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=1426; METHOD=FAE.
KW Amphibian skin.
SQ SEQUENCE 13 AA; 1457 MW; 290CA148C384B5A7 CRC64;

Query Match 18.8%; Score 18; DB 1; Length 13;
Best Local Similarity 33.3%; Pred. No. 3e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FEHCNFNDV 9
Db 1 FQFVNPDSI 9

RESULT 23
LECB_PSOSC STANDARD; PRT; 14 AA.
ID LECB_PSOSC
AC P22584:
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE BASIC LECTIN B1 (FRAGMENT).
OS Psophocarpus scandens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia; Euphorbiaceae; Papilionoideae; Phaseoleae;
OC Psophocarpus.
OC NCB1_TaxID=3690;
RN [1]
RP SEQUENCE.
RX MEDLINE=93126144; PubMed=1480511;
RA Strangler J., Hilbich C., Burdzik S., Keller R.;
RT Orckokin: a novel myotropic peptide from the nervous system of the
crayfish, Orconectes limosus.";
RL Peptides 13:859-864(1992).
CC -1- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS.
CC -1- TISSUE SPECIFICITY: ABDOMINAL NERVE CORD AND HINDGUT.
KW Neuropeptide.

```

CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.  
 CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL  
 CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
 DR PIR: PA0007; PA0007.  
 DR HSSP: 024313; 1MBL.  
 KW Lectin: Glycoprotein.  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1732 MW; D80ACE43B487C549 CRC64;

Query Match 18.8%; Score 18; DB 1; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 5 NFN 7  
 DB 6 NFN 8

RESULT 24  
 LEC1\_PSOSC STANDARD; PRT; 15 AA.  
 ID LEC1\_PSOSC  
 AC P22582;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ACIDIC LECTIN A1 (FRAGMENT).  
 OS Psophocarpus scandens.  
 CC Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;  
 CC Spermatophyta; Magnoliophyta: eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 OC Psophocarpus.  
 OX NCBI\_TaxID=3890;  
 RN [1]  
 RN SEQUENCE.  
 RP TISSUE=Seed;  
 RA Kott A.A.;  
 RT "Isolation and characterization of the lectins from the seeds of  
 RL Psophocarpus scandens.";  
 RL Phytochemistry 27:2847-2855(1988).  
 CC -1- SUBUNIT: DIMER. THE SUBUNITS SHOW APPARENT MW HETEROGENEITY  
 CC (32000-35000 MW), WHICH MAY RESULT FROM DIFFERENT CARBOHYDRATE  
 CC CONTENT, AA SEQUENCE, OR POLYPEPTIDE LENGTH.  
 CC -1- PTM: CONTAINS 5-6% CARBOHYDRATE.  
 CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS ACIDIC LECTINS IN N-TERMINAL  
 CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
 DR PIR: PA0005; PA0005.  
 KW Lectin: Glycoprotein.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1743 MW; D6EB9FE9C6B0254D CRC64;

Query Match 18.8%; Score 18; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 5 NFN 7  
 DB 7 NFN 9

RESULT 25  
 LEC2\_PSOSC STANDARD; PRT; 15 AA.  
 ID LEC2\_PSOSC  
 AC P22585;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE BASIC LECTIN B2 (FRAGMENT).  
 OS Psophocarpus scandens.  
 CC Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;  
 CC Spermatophyta; Magnoliophyta: eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

OC Psophocarpus.  
 OX NCBI\_TaxID=3890;  
 RN [1]  
 RN SEQUENCE.  
 RP TISSUE=Seed;  
 RA Kott A.A.;  
 RT "Isolation and characterization of the lectins from the seeds of  
 RL Psophocarpus scandens.";  
 RL Phytochemistry 27:2847-2855(1988).  
 CC -1- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF  
 CC ABOUT 32000 APPARENT MW.  
 CC -1- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE  
 CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.  
 CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL  
 CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
 DR PIR: PA0008; PA0008.  
 KW Lectin: Glycoprotein.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1847 MW; D194CE400C832796 CRC64;

Query Match 18.8%; Score 18; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 5 NFN 7  
 DB 7 NFN 9

Search completed: April 2, 2002, 09:27:44  
 Job time: 183 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:24:11 ; Search time 22.03 Seconds  
(without alignments)  
112.875 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTTRELENE 17

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
Total number of hits satisfying chosen parameters: 3442

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :  
1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID     | Description        |
|------------|-------|-------------|--------|--------|--------------------|
| 1          | 25.0  | 15          | 13     | Q9PRW3 | Q9PRW3 crocalus at |
| 2          | 25.0  | 15          | 13     | Q9PRW2 | Q9PRW2 crocalus at |
| 3          | 24.0  | 9           | 4      | O14715 | O14715 homo sapien |
| 4          | 24.0  | 15          | 4      | P78482 | P78482 homo sapien |
| 5          | 25.0  | 16          | 10     | Q9AXW2 | Q9AXW2 brassica na |
| 6          | 24.0  | 16          | 10     | Q9AXV9 | Q9AXV9 silene aega |
| 7          | 23.0  | 12          | 10     | Q9FSA9 | Q9FSA9 silene aega |
| 8          | 23.0  | 13          | 11     | P97622 | P97622 rattus norv |
| 9          | 24.0  | 14          | 2      | O55326 | O55326 synecococc  |
| 10         | 22.9  | 9           | 4      | O95574 | O95574 homo sapien |
| 11         | 22.9  | 16          | 7      | O77922 | O77922 oreochromis |
| 12         | 22.9  | 17          | 12     | Q9GEX2 | Q9GEX2 human immun |
| 13         | 21.9  | 9           | 5      | O96417 | O96417 drosophila  |
| 14         | 21.9  | 10          | 6      | Q9TQV4 | Q9TQV4 equus cabal |
| 15         | 21.9  | 10          | 10     | Q9FS93 | Q9FS93 silene pent |
| 16         | 21.9  | 12          | 6      | Q9TRU1 | Q9TRU1 bos taurus  |
| 17         | 21.9  | 12          | 10     | P82437 | P82437 nicotiana t |
| 18         | 21.9  | 13          | 4      | Q9UNV6 | Q9UNV6 homo sapien |
| 19         | 21.9  | 13          | 10     | Q9FSA8 | Q9FSA8 silene bac  |

|    |      |    |    |        |                     |
|----|------|----|----|--------|---------------------|
| 20 | 21.9 | 13 | 10 | Q9FS94 | Q9FS94 silene pent  |
| 21 | 21.9 | 14 | 10 | Q9FSB1 | Q9FSB1 silene aega  |
| 22 | 21.9 | 14 | 10 | Q9FS95 | Q9FS95 silene pent  |
| 23 | 21.9 | 14 | 10 | Q9FS91 | Q9FS91 silene sedo  |
| 24 | 21.9 | 15 | 10 | Q9FSB0 | Q9FSB0 silene aega  |
| 25 | 21.9 | 15 | 10 | Q9FSA5 | Q9FSA5 silene cryp  |
| 26 | 21.9 | 15 | 10 | Q9FSA0 | Q9FSA0 silene haus  |
| 27 | 21.9 | 15 | 10 | Q9FS99 | Q9FS99 silene inte  |
| 28 | 21.9 | 15 | 10 | Q9FS98 | Q9FS98 silene laco  |
| 29 | 21.9 | 15 | 10 | Q9FS80 | Q9FS80 silene zawa  |
| 30 | 21.9 | 15 | 10 | Q9FE15 | Q9FE15 silene sedo  |
| 31 | 21.9 | 15 | 10 | Q9FE14 | Q9FE14 silene pent  |
| 32 | 21.9 | 15 | 10 | Q9FE13 | Q9FE13 silene aega  |
| 33 | 21.9 | 16 | 4  | Q9BQF2 | Q9BQF2 homo sapien  |
| 34 | 21.9 | 16 | 11 | Q63039 | Q63039 rattus norv  |
| 35 | 21.9 | 16 | 12 | Q66737 | Q66737 equine infe  |
| 36 | 20.8 | 9  | 2  | Q48686 | Q48686 lactococcus  |
| 37 | 20.8 | 11 | 12 | Q66874 | Q66874 feline call  |
| 38 | 20.8 | 11 | 12 | Q66877 | Q66877 feline call  |
| 39 | 20.8 | 13 | 11 | Q54809 | Q54809 mus musculu  |
| 40 | 20.8 | 13 | 12 | Q86129 | Q86129 vesicular s  |
| 41 | 20.8 | 15 | 8  | Q95952 | Q95952 sauromalus   |
| 42 | 20.8 | 15 | 10 | Q958P1 | Q958P1 raphanus sa  |
| 43 | 20.8 | 15 | 11 | Q9QYB7 | Q9QYB7 rattus sp.   |
| 44 | 20.8 | 15 | 11 | Q9QY00 | Q9QY00 rattus sp.   |
| 45 | 20.8 | 16 | 5  | Q9TWS8 | Q9TWS8 lamellibrac  |
| 46 | 19.8 | 8  | 12 | Q84271 | Q84271 human papil  |
| 47 | 19.8 | 10 | 5  | Q9TWS1 | Q9TWS1 dermatophag  |
| 48 | 19.8 | 12 | 2  | Q9X645 | Q9X645 unidentified |
| 49 | 19.8 | 12 | 2  | Q9X643 | Q9X643 klebsiella   |
| 50 | 19.8 | 12 | 2  | Q9X640 | Q9X640 citrobacter  |

## ALIGNMENTS

RESULT 1  
Q9PRW3 PRELIMINARY; PRT; 15 AA.  
AC Q9PRW3  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ALPHA-FIBRINOGENASE ISOFORM A2 (FRAGMENT).  
OS Crocalus atrox (Western diamondback rattlesnake).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Crocodylia; Serpentes; Colubroidea;  
OC Viperidae; Crotalinae; Crotalus.  
OX NCBI\_TaxID=8730;  
RN [1]  
RP  
RX MEDLINE=94296418; Pubmed=8024586;  
RA Hung C.C., Chlou S.H.;  
RT "Isolation of multiple isoforms of alpha-fibrinogenase from the  
RT Western diamondback rattlesnake, Crotalus atrox: N-terminal sequence  
RT homology with anurod, an antithrombotic agent from Malaysian viper.";  
RT Biochem. Biophys. Res. Commun. 201:1414-1423(1994).  
SQ SEQUENCE 15 AA; 1640 MW; 03EPE10227CA12DA CRC64;

Query Match 26.0%; Score 25; DB 13; Length 15;  
Best Local Similarity 42.9%; Pred. No. 1.3e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 FEHCNFND 8  
Db 5 DECNINE 11

RESULT 2  
Q9PRW2 PRELIMINARY; PRT; 15 AA.  
ID Q9PRW2  
AC Q9PRW2;

DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE ALPHA-FIBRINOGENASE ISOFORM A3 (FRAGMENT).  
 OS Crocalus atrox (Western diamondback rattlesnake).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 CC Viperidae; Crotalinae; Crotalus.  
 NC NCBL\_TaxID=8730;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94296418; PubMed=8024586;  
 RA Hung C.C., Chiu S.H.;  
 RT "Isolation of multiple isoforms of alpha-fibrinogenase from the  
 RT Western diamondback rattlesnake, *Crocalus atrox*: N-terminal sequence  
 RT homology with anurod, an antithrombotic agent from Malayan viper.";  
 RL Biochem. Biophys. Res. Commun. 201:1414-1423(1994).  
 SQ SEQUENCE 15 AA; 1656 MW; 03EBF10227D52FDA CRC64;

Query Match  
 Best Local Similarity 26.0%; Score 25; DB 13; Length 15;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 EHCNFD 8  
 DB 5 DECINE 11

RESULT 3  
 ID 014715 PRELIMINARY; PRT; 9 AA.  
 AC 014715;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE KERATIN 14 (FRAGMENT).  
 GN KRT14.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92005680; PubMed=1717157;  
 RA Coulombe P., Hutton M., Letai A., Hebert A., Paller A., Fuchs E.;  
 RT "Point mutations in human keratin 14 genes of epidermolysis bullosa  
 RT simplex patients: genetic and functional analyses.";  
 RL Cell 66:1301-1311(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95072587; PubMed=7526926;  
 RA Yamaniishi K., Matsuki M., Konishi K., Yasuno H.;  
 RT "A novel mutation of Leu122 to Phe at a highly conserved hydrophobic  
 RT residue in the helix initiation motif of keratin 14 in epidermolysis  
 RT bullosa simplex.";  
 RL Hum. Mol. Genet. 3:1171-1172(1994).  
 DR EMBL; D28807; BAA05967.1; -.  
 KW Keratin.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 9 AA; 1138 MW; BE300AA449C456D6 CRC64;

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 4; Length 9;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 NFND 8  
 DB 4 NFND 7

RESULT 4  
 ID P78482 PRELIMINARY; PRT; 15 AA.  
 AC P78482;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE APOLOPROTEIN B48 (FRAGMENT).  
 GN APOB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 37-51 FROM N.A.  
 RX MEDLINE=88018019; PubMed=3659919;  
 RA Chen S.H., Habib G., Yang C.Y., Gu Z.W., Lee B.R., Weng S.A.,  
 RA Silberman S.R., Cai S.J., Deslypere J.P., Rosseneu M., Gotto A.M. Jr.,  
 RT "Apolipoprotein B-48 is the product of a messenger RNA with an organ-  
 RT specific in-frame stop codon.";  
 RL Science 238:363-366(1987).  
 DR EMBL; M18036; AAB51754.1; -.  
 KW Lipoprotein.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 15 AA; 1842 MW; 9172790C16559AEB CRC64;

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 4; Length 15;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 5 NFNDVTRLR 14  
 DB 2 NFNEKSLQ 11

RESULT 5  
 ID 09AXW2 PRELIMINARY; PRT; 16 AA.  
 AC 09AXW2;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DE PHYTOCHROME A (FRAGMENT).  
 GN BN-PRYA-1.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.  
 NC NCBL\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. STELLAR;  
 RA Foudmann M., Froger N., Brunel D.;  
 RT "Amplified consensus gene markers: Tools designing for a genetic map  
 RT of Arabidopsis-known-function genes in Brassica.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF229413; AAK00685.1; -.  
 KW Phytochrome.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 16 AA; 1791 MW; D390F9AD5DDFBFID CNC64;

Query Match  
 Best Local Similarity 25.0%; Score 24; DB 10; Length 16;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 DVTRLRN 16  
 DB 5 DIVORMLN 13

```

RESULT 6
O9AXV9 PRELIMINARY; PRT; 16 AA.
AC O9AXV9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE PHYCHROME A (FRAGMENT).
GN PHVA-BO-1.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. RAPID CYCLING;
RA Poutman M., Froger N., Brunel D.;
RT "Amplified consensus gene markers: Tools designing for a genetic map
of Arabidopsis-known-function genes in Brassica."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF229417; AAK00689.1; -.
FT NON_TER 1 1
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1819 MW; 04C0F9AD5DFBF0F CRC64;

```

```

Query Match 25.0%; Score 24; DB 10; Length 16;
Best Local Similarity 44.4%; Pred. No. 2e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 8 DVTTRLREN 16
Db 5 DIVORLEN 13

RESULT 7
O9FSA9 PRELIMINARY; PRT; 12 AA.
AC O9FSA9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene aegaea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49732;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296130; CAC13014.1; -.
FT NON_TER 1 1
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1304 MW; 83269695B441B2CA CRC64;

```

```

Query Match 24.0%; Score 23; DB 10; Length 12;
Best Local Similarity 55.6%; Pred. No. 2.2e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 6 FNDVTRLR 14
Db 3 FTDVTISAR 11

RESULT 8
P97622

```

```

ID P97622 PRELIMINARY; PRT; 13 AA.
AC P97622;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DE FERRITIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR-KYOTO; TISSUE=AORTIC SMOOTH MUSCLE;
RA Adams L.A., Werny I., Schwartz S.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U75408; AAB19110.1; -.
DR HSSP: P02791; IAEW.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1557 MW; 72A2122AF06D56C4 CRC64;

```

```

Query Match 24.0%; Score 23; DB 11; Length 13;
Best Local Similarity 37.5%; Pred. No. 2.4e+03;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 9 VTTTRLREN 16
Db 1 MTSQIRON 8

RESULT 9
O55326 PRELIMINARY; PRT; 14 AA.
AC O55326;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE PUTATIVE ORF1 (FRAGMENT).
OS Synechococcus sp.
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1131;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92201692; Pubmed=1551590;
RA Rhiel E., Stirewalt V.L., Gasparich G.E., Bryant D.A.;
RT "The psac genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa:
cloning and sequence analysis."
RL Gene 112:123-128(1992).
DR EMBL: M86238; AAA27351.1; -.
FT NON_TER 1 1
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1590 MW; 9D5226HDCB98BCA CRC64;

```

```

Query Match 24.0%; Score 23; DB 2; Length 14;
Best Local Similarity 71.4%; Pred. No. 2.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 8 DVTTRLR 14
Db 7 DVTGRLQ 13

RESULT 10
O95574 PRELIMINARY; PRT; 9 AA.
AC O95574;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE SERUM ALBUMIN (FRAGMENT).
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL LIVER;  
 RX MEDLINE=82081882; PubMed=6171778;  
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wilson K.L.,  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 RT coli.";  
 RL Nucleic Acids Res. 9:6103-6114(1981).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL LIVER;  
 RX MEDLINE=82105994; PubMed=6275391;  
 RA Dugaiczky A., Law S.W., Dennison O.E.,  
 RT "Nucleotide sequence and the encoded amino acids of human serum  
 RT albumin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL LIVER;  
 RX MEDLINE=85155492; PubMed=6085063;  
 RA Urano Y., Sakai M., Watanabe K., Tamaoki T.,  
 RT "Tandem arrangement of the albumin and alpha-fetoprotein genes in the  
 RT human genome.";  
 RL Gene 32:255-261(1984).  
 DR EMBL: L00133; AAC93594.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 1186 MW; 92340B0400440681 CRC64;

Query Match 22.9%; Score 22; DB 4; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+05;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Oy 10 TTRLRNE 17  
 | : : : :  
 Db 2 TMRIRERK 9

RESULT 11  
 ID 077922 PRELIMINARY; PRT; 16 AA.  
 AC 077922;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE MHC CLASS II B LOCUS 16 (FRAGMENT).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorphi; Acanthopterygii; Perciformes; Labroidel;  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98315113; PubMed=9649539;  
 RA Mataga-Trillo E., McAndrew B., Vincok V., Zaleska-Rutczynska Z.,  
 RA Stelmann H., Figueroa F., Klein J.,  
 RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
 RT class II B loci.";  
 RL Genetics 149:1527-1547(1998).  
 DR EMBL: AF050035; AAC41374.1; -.  
 FT MHC.  
 KW NON\_TER 1  
 SQ SEQUENCE 16 AA; 1893 MW; A56619E46B4B5F2D CRC64;

Query Match 22.9%; Score 22; DB 7; Length 16;  
 Best Local Similarity 75.0%; Pred. No. 4.2e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 4 CNFN 7  
 | : : :  
 Db 10 CDFN 13  
 RESULT 12  
 ID 090EX2 PRELIMINARY; PRT; 17 AA.  
 AC 090EX2;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin H.J., Siwak E.B., Hollinger F.B.,  
 RT "Mutation rate of human immunodeficiency virus type 1 genomic RNA  
 RT deduced from long term culture of its biological clones.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF178668; AAF04374.1; -.  
 KW Envelope protein.  
 FT NON\_TER 1  
 FT NON\_TER 17  
 SQ SEQUENCE 17 AA; 1897 MW; 3346BDC9A99CE117 CRC64;

Query Match 22.9%; Score 22; DB 12; Length 17;  
 Best Local Similarity 35.7%; Pred. No. 4.5e+03;  
 Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 Oy 4 CNFNVTTRLRNE 17  
 | : : : :  
 Db 2 CTFLNVTTRISTEE 15

RESULT 13  
 ID 096417 PRELIMINARY; PRT; 9 AA.  
 AC 096417;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DE SXL E1 FORM (FRAGMENT).  
 GN SXL.  
 OS Drosophila virilis (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7244;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98337843; PubMed=9671597;  
 RA Erickson J.W., Cline T.W.,  
 RA "Key aspects of the primary sex determination mechanism are conserved  
 RT across the genus Drosophila.";  
 RL Development 125:3259-3268(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bell M., Cline T.W.,  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF046045; AAC97605.1; -.  
 DR FlyBase: FBgn0016470; Dvir\Sxl.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1089 MW; 9A8BD1AA9C449CA CRC64;

Query Match 21.9%; Score 21; DB 5; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 4.7e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;



OY 5 NENDVT 10  
 11: 11  
 DB 4 NEDVT 9

# RESULT 14

O9TOV4 PRELIMINARY: PRT: 10 AA.  
 AC O9TOV4:  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE POP-VARIANT \* TRANSERRIN GENE, EXON 17 AND PARTIAL CDS (FRAGMENT).  
 OS Equus caballus (horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Giffard J.M., Brandon R.B., Bell T.K.;  
 RT "Further identification of single nucleotide polymorphisms in the  
 RT equine transferrin gene."  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF185778; AAF05491.1; -;  
 DR EMBL; AF185770; AAF05483.1; -;  
 DR EMBL; AF185771; AAF05484.1; -;  
 DR EMBL; AF185772; AAF05485.1; -;  
 DR EMBL; AF185773; AAF05486.1; -;  
 DR EMBL; AF185774; AAF05487.1; -;  
 DR EMBL; AF185775; AAF05488.1; -;  
 DR EMBL; AF185776; AAF05489.1; -;  
 DR EMBL; AF185777; AAF05490.1; -;  
 FT NON\_TER 1  
 SO SEQUENCE 10 AA; 1188 MW; 6DF67DE9D1AEBDB CRC64;

Query Match 21.9%; Score 21; DB 6; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3.9e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 EHCNENDV 9  
 11: 11  
 DB 3 EACTFHVR 10

# RESULT 15

O9FS93 PRELIMINARY: PRT: 10 AA.  
 AC O9FS93:  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE RNA POLYMERASE II (FRAGMENT).  
 GN RPB2.  
 OS Silene pentellica.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.  
 OX NCBI\_TaxID=49735;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Popp M., Oxelman B.;  
 RT "Inferring the history of the polyploid silene aegaea  
 RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data."  
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; A0296133; CAC13025.1; -;  
 FT NON\_TER 1  
 SO SEQUENCE 10 AA; 1108 MW; CF1AB6D1B2CAB1A9 CRC64;

Query Match 21.9%; Score 21; DB 10; Length 10;

Best Local Similarity 80.0%; Pred. No. 3.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 FNDVT 10  
 11: 11  
 DB 3 FNDVT 7

# RESULT 16

O9TRU1 PRELIMINARY: PRT: 12 AA.  
 AC O9TRU1:  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE 43 KDA CYANOGEN BROMIDE FRAGMENT PEAK 7.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92132498; PubMed=1734497;  
 RA Velby O.P., Sletten K., Husby G., Nordstoga K.;  
 RT "Amino acid sequence analyses of non-AA proteins from amyloid fibrils  
 RT of bovine kidney."  
 RL Scand. J. Immunol. 35:63-69(1992).  
 SO SEQUENCE 12 AA; 1257 MW; 0D5C94FDE9B76AAA CRC64;

Query Match 21.9%; Score 21; DB 6; Length 12;  
 Best Local Similarity 80.0%; Pred. No. 4.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 FNDVT 10  
 11: 11  
 DB 4 FNDPT 8

# RESULT 17

P82437 PRELIMINARY: PRT: 12 AA.  
 ID P82437:  
 DT 01-JUN-2000 (TREMblrel. 14, Created)  
 DT 01-JUN-2000 (TREMblrel. 14, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE 26 KDA CELL WALL PROTEIN (FRAGMENT).  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-CV. PETT HAVANA;  
 RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,  
 RA Wojtaszek P., Bolwell G.P.;  
 RT "Proteomic study of secondary cell wall proteins from transformed  
 RT tobacco culture."  
 RL Planta 0:0-0(2000).  
 CC -I- SUBCELLULAR LOCATION: CELL WALL.  
 CC -I- TISSUE SPECIFICITY: XYLEM.  
 KM Cell wall.  
 FT NON\_TER 12  
 SO SEQUENCE 12 AA; 1452 MW; 92B6CDD0F3A45AF1 CRC64;

Query Match 21.9%; Score 21; DB 10; Length 12;  
 Best Local Similarity 80.0%; Pred. No. 4.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 NDVT 11

Db 7 NDVXT 11

```

RESULT 18
ID 09UNV6 PRELIMINARY: PRT: 13 AA.
AC 09UNV6:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE INOSINE MONOPHOSPHATASE 2 (FRAGMENT).
GN IMP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97463449; PubMed=9322233;
RA Yoshikawa T., Turner G., Esterling L.E., Sanders A.R.,
RA Deleter-Wadleigh S.D.;
RT "A novel human myo-inositol monophosphatase gene, IMP.18p, maps to a
RT susceptibility region for bipolar disorder.";
RL Mol. Psych. 2:393-397(1997).
DR EMBL: AF085628; AAD22141.1;
DR EMBL: AF085627; AAD22141.1; JOINED.
DR HSP: P29218; IIMB.
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1491 MW; E8154075786DD401 CRC64;

```

Query Match 21.9%; Score 21; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 5e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 CNF 6  
 111  
 2 CNF 4

```

RESULT 19
ID 09FSB8 PRELIMINARY: PRT: 13 AA.
AC 09FSB8:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene bacillifera (Berry catchfly) (Cucubalus bacclifer).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=34818;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296139; CAC13015.1;
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1437 MW; D931E2CF1AB6D1B2 CRC64;

```

Query Match 21.9%; Score 21; DB 10; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 6 FNDVT 10

Db 3 FTDVT 7

```

RESULT 20
ID 09FS94 PRELIMINARY: PRT: 13 AA.
AC 09FS94:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene pentellica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49735;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296132; CAC13024.1;
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1437 MW; D931E2CF1AB6D1B2 CRC64;

```

Query Match 21.9%; Score 21; DB 10; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 6 FNDVT 10  
 111  
 3 FTDVT 7

```

RESULT 21
ID 09FSB1 PRELIMINARY: PRT: 14 AA.
AC 09FSB1:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene aegaea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49732;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ296122; CAC13007.1;
FT NON_TER 1
FT NON_TER 14
SQ SEQUENCE 14 AA; 1508 MW; 9B0931E2CF1AB6D1 CRC64;

```

Query Match 21.9%; Score 21; DB 10; Length 14;  
 Best Local Similarity 80.0%; Pred. No. 5.4e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 6 FNDVT 10  
 111  
 3 FTDVT 7

```

RESULT 22
09FS95 PRELIMINARY; PRT; 14 AA.
AC 09FS95:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene pentellica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49735;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296131; CAC13023.1; -.
FT NON_TER 1 1
SQ SEQUENCE 14 AA; 1508 MW; 9B0931E2CF1AB6D1 CRC64;

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 14;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 FNDVT 10
DB 3 FTDVT 7

```

```

RESULT 23
09FS91 PRELIMINARY; PRT; 14 AA.
AC 09FS91:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene sedoides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=39790;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296105; CAC13028.1; -.
FT NON_TER 1 1
SQ SEQUENCE 14 AA; 1508 MW; 9B0931E2CF1AB6D1 CRC64;

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 14;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 FNDVT 10
DB 3 FTDVT 7

```

RESULT 24  
09FS80

```

ID 09FS80 PRELIMINARY; PRT; 15 AA.
AC 09FS80:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene aegaea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=49732;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296124; CAC13009.1; -.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1616 MW; C96B0931E2CF0AF6 CRC64;

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 15;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 FNDVT 10
DB 3 FTDVT 7

```

```

RESULT 25
09FSAS PRELIMINARY; PRT; 15 AA.
AC 09FSAS:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE RNA POLYMERASE II (FRAGMENT).
GN RPB2.
OS Silene cryptoneura.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX NCBI_TaxID=39877;
RN [1]
RP SEQUENCE FROM N.A.
RA Popp M., Oxelman B.;
RT "Inferring the history of the polyploid Silene aegaea
(RL (Caryophyllaceae) using nuclear and chloroplast DNA sequence data.";
DR EMBL; AJ296138; CAC13016.1; -.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1621 MW; C96B0931E2CF1AB6 CRC64;

```

```

Query Match
Best Local Similarity 21.9%; Score 21; DB 10; Length 15;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 FNDVT 10
DB 3 FTDVT 7

```

Search completed: April 2, 2002, 09:27:28  
Job time: 197 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:22:51 ; Search time 23.72 Seconds

(without alignments)  
53.088 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58

Perfect score: 96  
Sequence: 1 FEHCNFNDVTRRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 183508

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 50 summaries

Database :

A\_Geneseq\_1101:\*

1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:\*

2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:\*

3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:\*

4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:\*

5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:\*

6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:\*

7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:\*

8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:\*

9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:\*

10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:\*

11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:\*

12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:\*

13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:\*

14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:\*

15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:\*

16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:\*

17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:\*

18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:\*

19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:\*

20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT:\*

21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:\*

22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 38    | 39.6        | 16     | 22 | AA67300     |
| 2          | 32.5  | 33.9        | 16     | 21 | AA58596     |
| 3          | 32    | 33.3        | 15     | 16 | AA83971     |
| 4          | 32    | 33.3        | 16     | 16 | AA70779     |
| 5          | 32    | 33.3        | 17     | 18 | AA30024     |
| 6          | 32    | 33.3        | 17     | 19 | AA66142     |
| 7          | 30    | 31.2        | 8      | 21 | AA61736     |
| 8          | 30    | 31.2        | 9      | 18 | AA43797     |
| 9          | 30    | 31.2        | 9      | 21 | AA61739     |
| 10         | 30    | 31.2        | 12     | 22 | AA97334     |
| 11         | 30    | 31.2        | 15     | 19 | AA41168     |

|    |      |      |    |    |          |                     |
|----|------|------|----|----|----------|---------------------|
| 12 | 30   | 31.2 | 16 | 19 | AA41167  | Metal-regulated tr  |
| 13 | 29   | 30.2 | 9  | 17 | AA49475  | Human leucocyte an  |
| 14 | 29   | 30.2 | 14 | 16 | AA871732 | Merosin major subu  |
| 15 | 29   | 30.2 | 16 | 21 | AA36539  | Human serum albumi  |
| 16 | 29   | 30.2 | 16 | 21 | AA36546  | Human serum albumi  |
| 17 | 28   | 29.2 | 6  | 3  | AA20158  | Hexapeptide #5 vit  |
| 18 | 28   | 29.2 | 9  | 19 | AA720456 | Human microtubule   |
| 19 | 28   | 29.2 | 10 | 22 | AA43078  | Mycoplasma genital  |
| 20 | 28   | 29.2 | 13 | 21 | AA57800  | TRAM-interaction m  |
| 21 | 28   | 29.2 | 15 | 12 | AA811465 | Lepidoptera allost  |
| 22 | 27.5 | 28.6 | 10 | 18 | AA432738 | Human platelet gly  |
| 23 | 27   | 28.1 | 12 | 20 | AA933135 | Human growth hormo  |
| 24 | 27   | 28.1 | 12 | 21 | AA823282 | Hyaluronic acid-bl  |
| 25 | 27   | 28.1 | 13 | 9  | AA81812  | Sequence of peptid  |
| 26 | 27   | 28.1 | 13 | 9  | AA81812  | Sequence of peptid  |
| 27 | 27   | 28.1 | 13 | 13 | AA822784 | HIV-2 env protein   |
| 28 | 27   | 28.1 | 14 | 22 | AA801700 | Human gene 3 encod  |
| 29 | 27   | 28.1 | 15 | 22 | AA84963  | Apoptin-associated  |
| 30 | 27   | 28.1 | 16 | 21 | AA98301  | Alpha D peptide de  |
| 31 | 27   | 28.1 | 17 | 21 | AA98399  | Alpha D peptide de  |
| 32 | 26.5 | 27.6 | 15 | 17 | AAW07941 | gp120 peptide HPI0  |
| 33 | 26   | 27.1 | 11 | 21 | AA815256 | Specific factor VI  |
| 34 | 26   | 27.1 | 11 | 21 | AA815259 | Specific factor VI  |
| 35 | 26   | 27.1 | 12 | 14 | AA835259 | D32.39 antibody is  |
| 36 | 26   | 27.1 | 12 | 15 | AA86727  | Random peptide #7   |
| 37 | 26   | 27.1 | 12 | 17 | AA81475  | D32.39 monoclonal   |
| 38 | 26   | 27.1 | 12 | 18 | AAW25257 | Antibody D32.39 ep  |
| 39 | 26   | 27.1 | 13 | 22 | AA889315 | HIV gp120 protein   |
| 40 | 26   | 27.1 | 13 | 22 | AA848698 | Mouse IL-3 recepto  |
| 41 | 26   | 27.1 | 14 | 15 | AA869361 | Stearoyl-VLP(1-14)  |
| 42 | 26   | 27.1 | 14 | 15 | AA94563  | Vasodilative intest |
| 43 | 26   | 27.1 | 15 | 22 | AA889312 | HIV gp120 protein   |
| 44 | 26   | 27.1 | 16 | 21 | AA979369 | Human CXCR3 immuno  |
| 45 | 26   | 27.1 | 16 | 21 | AA956017 | Mucor hiemalis end  |
| 46 | 25.5 | 26.6 | 12 | 17 | AAV24486 | Actinomodura sp. D  |
| 47 | 25.5 | 26.6 | 12 | 17 | AA94887  | Actinomodura flexu  |
| 48 | 25   | 26.0 | 6  | 18 | AAW03262 | Viral integrase in  |
| 49 | 25   | 26.0 | 7  | 21 | AA61734  | Cadherin-8 cell ad  |
| 50 | 25   | 26.0 | 10 | 18 | AAW36662 | Thrombopoietin rec  |

#### ALIGNMENTS

RESULT 1

AA67300

ID AAB67300 standard; Peptide: 16 AA.

XX AAB67300;

AC

DT 20-APR-2001 (first entry)

XX

DE Cytokine peptide #6.

XX

XX

KW Prosaposin; cell death; Bcl-XL; brain; cardiac muscle.

XX

OS Homo sapiens.

XX

PN EPI072609-AA2.

PD

PD 31-JAN-2001.

XX

PF 30-JUN-2000; 2000EP-0305504.

PR 30-JUN-1999; 99JP-0185155.

XX

PA (SAKA/) SAKANAKA M.

PA (TANA/) TANAKA J.

PA (SATO/) SATO K.

XX

PI Sakanaka M, Tanaka J, Sato K, Morita F, Sadamoto Y;

DR WPI; 2001-204263/21.

XX Use of prosaposin-related peptides or derivatives as cytoprotective  
 PT agents, for suppressing apoptosis or apoptosis-like cell death -  
 XX  
 PS Disclosure; Page 12; 41pp; English.  
 CC  
 CC The present invention relates to use of a prosaposin-related peptide  
 CC or derivative, in the production of a medicament for use in  
 CC preventing or delaying cell death, or in promoting the expression  
 CC of cell death supporting gene product Bcl-XL. The invention is  
 CC useful for preventing the death of cells e.g. brain cells,  
 CC neurons and cardiac muscle cells, in vitro or ex vivo.  
 XX  
 SQ Sequence 16 AA;

Query Match 39.6%; Score 38; DB 22; Length 16;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EHCNNDVTR 11  
 ||| : : ||  
 Db 2 EHCNNDVTR 11

RESULT 2  
 AAY58596  
 ID AAY58596 standard; peptide: 16 AA.  
 XX  
 AC AAY58596;  
 XX  
 DT 10-APR-2000 (first entry)  
 XX

XX Toxoplasma gondii antigen SAG1 epitope #8.  
 XX  
 KW SAG1: antigen; toxoplasmosis; subunit vaccine; Pichia pastoris;  
 KM epitope; antibody.  
 XX

OS Toxoplasma gondii.  
 XX  
 PN WO966043-A1.  
 XX  
 PD 23-DEC-1999.  
 XX

PF 08-JUN-1999; 99WO-EP03957.  
 XX  
 PR 12-JUN-1998; 98GB-0012773.  
 PR 15-APR-1999; 98GB-0008564.  
 XX

XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX

PI Biemann R, Bollen A, Haumont M;  
 PT  
 DR WPI; 2000-106101/09.  
 XX

XX Method for production of toxoplasma antigen SAG1 for use in vaccines -

XX Example 1; Page 14; 47pp; English.  
 XX

XX The invention relates to the recombinant production of the toxoplasma  
 CC antigen SAG1 or a fragment thereof in the yeast *Pichia pastoris*. SAG1  
 CC is the major surface antigen of *Toxoplasma gondii*, an obligate  
 CC intracellular protozoan parasite responsible for toxoplasmosis in  
 CC mammals, including humans. SAG1 is anchored to the plasma membrane of  
 CC *T. gondii* via a GPI (91ycosylphosphatidylinositol) anchor, which is  
 CC attached to the C-terminal anchor region (residues 308-336). The  
 CC invention provides a truncated SAG1 protein, which lacks the anchor  
 CC region and comprises amino acids 48-307, and a SAG protein N-terminally  
 CC fused to a yeast secretion signal peptide. The SAG protein and its  
 CC fragments can be used in the manufacture of a subunit vaccine for the  
 CC prevention or treatment of toxoplasmosis in mammals. A live attenuated  
 CC Toxoplasma vaccine is available, but cannot be administered to humans  
 CC due to the risk of reversion of the attenuated strain to a virulent

CC form. Prior art recombinant expression of SAG1 has been attempted in  
 CC *Escherichia coli*, mammalian cells or *S. cerevisiae*, but has been  
 CC associated with problems such as misfolding and insolubility. In  
 CC of correctly folded SAG1, or heterogeneous protein production. In  
 CC addition, purification of the native protein from tachyzoites is  
 CC difficult and time-consuming, due to the GPI anchor. Sequences  
 CC AAY58595-Y58596 represent peptide epitopes of *T. gondii* SAG1, which  
 CC were used to raise polyclonal antibodies in an exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 16 AA;

Query Match 33.9%; Score 32.5; DB 21; Length 16;  
 Best Local Similarity 43.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 2; Mismatches 4; Indels 3; Gaps 1;  
 QY 4 CN---FNDVTRLRN 16  
 || | : : ||  
 Db 1 CNEKFKDLPKLTN 16

RESULT 3  
 AAR83971  
 ID AAR83971 standard; peptide: 15 AA.  
 XX  
 AC AAR83971;  
 XX  
 DT 21-MAY-1996 (first entry)  
 XX

XX Virus infection factor N-terminal peptide from silkworm body fluid.  
 XX  
 DE  
 XX  
 KW Amino terminal; silkworm; virus infection factor; promoter;  
 KM recombinant; high yield; body fluid.  
 XX

OS *Phloesamia cynthia* ricini.  
 XX  
 PN JP07252298-A.  
 XX  
 PD 03-OCT-1995.  
 XX

PE 18-NOV-1994; 94JP-0308468.  
 XX  
 PR 22-NOV-1993; 93JP-0314038.  
 XX

PA (KATA ) KATAKURA IND CO LTD.  
 XX

DR WPI; 1995-371176/48.  
 XX

XX Virus infection factor from silkworm body fluid - used to promote  
 PT infection of insect cells with protein-expressing recombinant virus  
 PT in culture medium  
 XX

PS Claim 6; Page 2; 11pp; Japanese.  
 XX

XX AAR83791 is an amino-terminal peptide of a virus infection factor  
 CC derived from heat-treated silkworm body fluid. The new virus  
 CC infection factor (VIF) is useful for promoting the infection of insect  
 CC cells with recombinant nuclear polyhedrosis virus. The virus having  
 CC been recombined with a useful protein-expressing gene. The insect  
 CC cell can then be cultured and the useful protein expressed at high  
 CC yields. The new VIF has a mol. wt. of approx. 15.2 kDa and an  
 CC isoelectric point of 9 or higher.  
 XX

SQ Sequence 15 AA;

Query Match 33.3%; Score 32; DB 16; Length 15;  
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 FNDVTRRL 13  
 || |||||

Db 2 fmvttl1 9

RESULT 4

ID AAR70779 standard; Peptide: 16 AA.

XX AAR70779;

XX 30-AUG-1995 (first entry)

XX

DE EPO neurotrophic peptide.

XX

XX Erythropoietin; EPO; cytokine; myelination; nervous system;

KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;

KW adrenal leukodystrophy; neuroblastoma.

XX

OS Homo sapiens.

XX WO9503821-A.

PN

XX 09-FEB-1995.

PD

XX 28-JUL-1994; 94MO-US08453.

PF

XX 30-JUL-1993; 93US-0100247.

PR

PR 21-APR-1994; 94US-0232513.

XX

PA (OBRI/) OBRIEN J S.

XX

PI Kishimoto Y, Obrien JS;

XX

DR WPI; 1995-082029/11.

XX

PT Stimulating neural cell out-growth and myelination - with

PT pro:saposin, saposin C or new neurotrophic peptide(s) from

PT cytokine(s), for treating nervous system diseases

XX

PS Disclosure; Page 36; 50pp; English.

XX

XX The peptide given in AAR70773, corresponding to amino acids 8-29 of

CC human saposin-C (AAR70784), promotes neurite outgrowth in vitro.

CC A consensus sequence was determined by comparing the peptide with

CC hematopoietic and neurotrophic cytokines, and neurotrophic peptides

CC (AAR70774-82) were identified in the AB loop of human ciliary

CC neurotrophic factor, interleukins-6, -2, -3 and -gamma,

CC erythropoietin and leukocyte inhibitory factor, and in helix C of human

CC interleukin-1-beta and oncostatin-M. Prosapoin (AAR70783)

CC and saposin-C also promoted nerve cell myelination ex vivo.

CC

XX

SQ Sequence 16 AA;

Query Match 33.3%; Score 32; DB 16; Length 16;

Best Local Similarity 50.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNFDVTT 11

|||:|:|

Db 2 ehcslnenit 11

RESULT 5

ID AAM30024 standard; peptide: 17 AA.

XX AAM30024;

XX 14-APR-1998 (first entry)

XX

DE Cytokine hepo AB loop derived peptide.

XX

KW Human; prosapoin; neural disorder; demyelination disorder;

KW neural cell death; inhibition; myelination; neurite outgrowth;

KW neuropathic pain; peripheral nerve disorder; neuroma; mononeuropathy;

KW polyneuropathy; dorsal root ganglia; spinal cord; brainstem;

KW thalamus; cytokine.

XX

OS Homo sapiens.

XX

XX WO9732895-A1.

PN

XX 12-SEP-1997.

PD

XX 05-MAR-1997; 97MO-US04143.

PF

XX 05-MAR-1996; 96US-0611307.

PR

XX (REGC ) UNIV CALIFORNIA.

PA

XX O'Brien JS;

PI

PI WPI; 1997-470538/43.

DR

XX

XX Prosapoin-derived peptide - useful for therapy of neural or

PT demyelination disorders in neural tissue

XX

PS Claim 16; Page 56; 69pp; English.

XX

XX The present sequence represents a specifically claimed cytokine derived

CC peptide for use in the following method of the invention. A method

CC has been developed of alleviating or preventing neuropathic pain in a

CC subject, comprising administering an effective amount of an active

CC fragment of prosapoin to the subject. The prosapoin-derived peptide

CC is useful for therapy of neural or demyelination disorders in neural

CC tissue. It can be used to stimulate neurite outgrowth, inhibit neural

CC cell death, promote myelination or inhibit demyelination. The method

CC is used to alleviate neuropathic pain resulting from a peripheral nerve

CC disorder, such as neuroma, nerve compression, crush or stretch and

CC incomplete nerve transection, mononeuropathy or polyneuropathy.

CC Alternatively the neuropathic pain results from a disorder of the

CC dorsal root ganglia, spinal cord, brainstem, thalamus or cortex.

CC

XX

SQ Sequence 17 AA;

Query Match 33.3%; Score 32; DB 18; Length 17;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNFDVTT 11

|||:|:|

Db 2 ehcslnenit 11

RESULT 6

ID AAM66142 standard; peptide: 17 AA.

XX AAM66142;

XX 17-NOV-1998 (first entry)

XX

DE Prosapoin receptor agonist #16.

XX

XX prosapoin; receptor agonist; neuropathic pain; neurite outgrowth;

KW neural cell death; nerve disorder; side effect.

XX

OS Synthetic.

OS Homo sapiens.

XX

XX WO9839357-A1.

PN

XX 11-SEP-1998.

PD

XX 11-SEP-1997; 97MO-US16062.

PF

XX 05-MAR-1997; 97MO-US04143.  
 PR (REGC ) UNIV CALIFORNIA.  
 PA  
 XX  
 PI O'Brien JS;  
 XX  
 DR WPI; 1998-495790/42.  
 XX  
 PT Use of prosaposin receptor agonists - for alleviating neuropathic  
 PT pain, inhibiting sensory or motor neuropathy, or inhibiting neural  
 XX cell death  
 PS  
 XX Claim 13; Page 19; 67pp; English.  
 CC  
 CC The invention relates to prosaposin receptor agonists. Also claimed  
 CC are: (1) methods for alleviating neuropathic pain or inhibiting the  
 CC onset of neuropathic pain, comprising administering a prosaposin receptor  
 CC agonist; (2) inhibiting sensory or motor neuropathy, comprising  
 CC contacting neuronal cells with a composition comprising a prosaposin  
 CC receptor agonist; (3) methods for stimulating neurite outgrowth,  
 CC inhibiting neural cell death, promoting myelination, or inhibiting  
 CC demyelination comprising contacting neuronal cells with a composition  
 CC comprising a prosaposin receptor agonist which has 14-50 amino acids and  
 CC may be used for treatment of neuropathic pain resulting from peripheral  
 CC nerve disorders (e.g. neuroma, nerve compression, nerve crush, nerve  
 CC stretch, incomplete nerve transection, mononeuropathy or polyneuropathy)  
 CC or disorders of dorsal root ganglia, the spinal cord, the brainstem, the  
 CC thalamus or the cortex. The receptor agonists do not cause undesirable  
 CC side effects. The present sequence represents a prosaposin receptor  
 CC agonist.  
 XX  
 SQ Sequence 17 AA;  
 DE  
 XX  
 Query Match 33.3%; Score 32; DB 19; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 OY 2 EHCNFNDVTT 11  
 III: I: I  
 Db 2 enclenmit 11  
 DE  
 XX  
 RESULT 7  
 AAY6I736  
 ID AAY6I736 standard; Peptide; 8 AA.  
 AC AAY6I736;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Cadherin-8 cell adhesion recognition cyclic peptide SEQ ID NO:1594.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; Li-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1..8  
 XX  
 PN \*WO9957149-A2.  
 XX  
 PD 11-NOV-1999.

XX 05-MAY-1999; 99MO-CA00363.  
 PR  
 XX  
 PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 PA  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 DR WPI; 2000-038791/03.  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 PS Claim 42; Page 176; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing neurotoxic disease,  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAY33183 to AAY33186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 8 AA;  
 DE  
 XX  
 Query Match 31.2%; Score 30; DB 21; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 4.3e+05;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 4 CNFNDVTT 10  
 I: I: I: I: I: I: I: I:  
 Db 1 cqnndvt 7  
 DE  
 XX  
 RESULT 8  
 AAM43797  
 ID AAM43797 standard; peptide; 9 AA.  
 XX  
 AC AAM43797;  
 XX  
 DT 20-APR-1998 (first entry)  
 XX  
 DE Specific human leukocyte antigen binding peptide #1.  
 XX  
 KW Human leukocyte antigen; HLA; vaccine; bacterial infection; viral;  
 KW cytotoxic T-cell; CTL; immunogenic peptide; cancer.  
 XX  
 OS Synthetic.  
 OS Haemophilus influenzae.  
 XX



PN WO9734617-A1.  
 XX  
 PD 25-SEP-1997.  
 XX  
 PF 21-MAR-1997; 97WO-US04451.  
 XX  
 PR 20-MAR-1997; 97US-0821739.  
 PR 21-MAR-1996; 96US-0013833.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Celis E, Grey HM, Kubo RT, Sette A;  
 XX  
 DR WPI: 1997-489250/45.  
 XX  
 PT Specific human leukocyte antigen binding peptide - used in vaccines  
 PT for the treatment and prevention of e.g. bacterial or viral  
 PT infection and cancer  
 PS  
 PS Claim 19; Page 37; 49pp; English.  
 XX  
 CC The present sequence represents a specific example of an immunogenic  
 CC peptide which was used in a new method of inducing a cytotoxic T cell  
 CC (CTL) response against a preselected antigen in a patient. The method  
 CC comprises contacting CTLs from the patient with the immunogenic peptide  
 CC (containing defined motifs) which binds one of the four HLA MHC products  
 CC HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation  
 CC constant (Kd) of less than 5 x 10<sup>-7</sup> M. Immunogens are viral, e.g. human  
 CC immunodeficiency virus type I (HIV-1), hepatitis B virus (HBV) and  
 CC hepatitis C virus (HCV) or cancer antigens and are used in vaccines for  
 CC the prevention and treatment of viral infection and cancer. The  
 CC immunogens may be administered to the patient as a nucleic acid encoding  
 CC the peptide (gene vaccine).  
 XX  
 SQ Sequence 9 AA:  
 OY 3 HCNPNDVT 10  
 Db 1 hsnlndt 8  
 MATCHES 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Query Match 31.2%; Score 30; DB 18; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 4.3e+05;  
 RESULT 9  
 ID AAY61739 standard; Peptide; 9 AA.  
 XX  
 AC AAY61739;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Cadherin-8 cell adhesion recognition cyclic peptide SEQ ID NO:1597.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW Ob-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; Li-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 1..9  
 PN WO957149-A2.

PD 11-NOV-1999.  
 XX  
 PF 05-MAY-1999; 99WO-CA00363.  
 XX  
 PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX  
 DR WPI: 2000-038791/03.  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 PS  
 PS Claim 42; Page 176; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MA's can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AA233183 to AA233186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 9 AA:  
 OY 4 CNEPDVT 10  
 Db 1 cgnldvt 7  
 MATCHES 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Query Match 31.2%; Score 30; DB 21; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 4.3e+05;  
 RESULT 10  
 ID AAB97334 standard; peptide; 12 AA.  
 XX  
 AC AAB97334;  
 XX  
 DT 13-AUG-2001 (first entry)  
 XX  
 DE Collagen IV alpha 3 domain epitope peptide #2.  
 XX  
 KW B cell; toxin; antigen specific; antibody mediated disease; virucide;  
 KW immunosuppressive; antiinflammatory; antiallergic; antidiabetic;  
 KW thyromimetic; antithyroid; vasotropic; cardiatic; antulcer;  
 KW neuroprotective; antineumatic; antiarthritic; dermatological;  
 KW ophthalmological; nephrotropic; allergy; autoimmune disorder;



PA (DART-) DARTMOUTH COLLEGE.  
 XX (MIND ) UNIV MINNESOTA.  
 XX  
 PI Elide DJ, Guerino ML;  
 XX  
 DR WPI: 1998-032230/03.  
 XX  
 PT Novel A. thaliana metal-regulated transporter proteins and related  
 PT DNA - for generating transgenic plants useful for removing  
 PT pollutants from soil or providing nutrient to patients suffering  
 PT from metal-deficiency disorders  
 XX  
 PS Claim 39; Page 105; 144pp; English.  
 XX  
 CC The metal regulated transporter epitope spans amino acids 162 through  
 CC 184 of IRT1 (AAV12763). The peptide was used as an antigen to raise  
 CC polyclonal antibodies. These antibodies recognize a protein of  
 CC approximately 33kDa which is only present in iron starved plants.  
 XX  
 SQ Sequence 16 AA;

Query Match 31.2%; Score 30; DB 19; Length 16;  
 Best Local Similarity 45.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTRLENE 17  
 ||| :|:|:  
 Db 3 ndvtrleped 13

RESULT 13  
 AAM49475  
 ID AAM49475 standard; peptide; 9 AA.

XX AAM49475;  
 XX  
 DT 05-JUN-1998 (first entry)  
 XX  
 DE Human leucocyte antigen DQ4 binding peptide #366.  
 XX  
 KW Human leucocyte antigen; HLA-DQ4; combinatorial library;  
 KW autoimmune disease; chronic articular rheumatism.  
 XX  
 OS Synthetic.  
 XX  
 PN JPO8151396-A.  
 PD 11-JUN-1996.  
 XX  
 PF 28-NOV-1994; 94JP-0292657.  
 XX  
 PR 28-NOV-1994; 94JP-0292657.  
 XX  
 PA (TEIJ ) TEIJIN LTD.  
 DR WPI: 1996-329479/33.  
 XX  
 PT HLA-binding oligopeptide and an immuno:regulator contg it - used in  
 PT the treatment of auto:immune disease  
 XX  
 PS Claim 4; Page 43; 61pp; Japanese.  
 XX  
 CC This peptide is an example of a peptide which binds to a human leucocyte  
 CC antigen HLA-DQ4 molecule. The peptide was isolated from a phagend  
 CC combinatorial library comprising the sequence AAV05953, by screening  
 CC with an HLA-DQ4 molecule. The peptide is used for the treatment of  
 CC autoimmune disease, or especially for treatment of viral diseases.  
 XX  
 SQ Sequence 9 AA;

Query Match 30.2%; Score 29; DB 17; Length 9;

Best Local Similarity 80.0%; Pred. No. 4.3e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 3 HCNFN 7  
 ||| |  
 Db 2 hcnfn 6

RESULT 14  
 AAR71732  
 ID AAR71732 standard; Peptide; 14 AA.

XX AAR71732;  
 XX  
 DT 01-MAY-1996 (first entry)  
 XX  
 DE Merosin major subunit derived peptide #2.

XX Human; 380-400 kD; merosin; major subunit; placenta; striated muscle;  
 KW peripheral nerve; trophoblast; Schwann cell neoplasm; 65 kD subunit;  
 KW 80 kD subunit; merosin polypeptide; merosin subunit; M chain;  
 KW laminin M chain; antigen; antibody; detection; tumour; malignancy;  
 KW neurite outgrowth; inhibitor; cell attachment.

XX Homo sapiens.  
 XX  
 PN MO9508628-A2.  
 PD 30-MAR-1995.  
 XX  
 PF 21-SEP-1994; 94MO-US10730.  
 XX  
 PR 22-SEP-1993; 93US-0125077.  
 XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.  
 XX  
 PI Engvall E, Lelivo I;  
 XX  
 DR WPI: 1995-139597/18.

XX  
 PT New merosin fragments, corresp. DNA and antibodies - for diagnosing  
 PT tumour malignancy, promoting or inhibiting neurite growth and  
 PT promoting cell attachment.  
 XX  
 PS Claim 6; Page 45; 65pp; English.

XX  
 CC The sequences given in AAR71731-32 represent peptides derived from  
 CC the human 380-400 kD merosin major subunit protein. Merosin is an  
 CC isoform of laminin and shows structural and sequence similarity to  
 CC the human laminin A chain. Mature human merosin is 30 amino acids  
 CC larger than the human laminin A chain. Similarly to all laminin chains,  
 CC the merosin protein has distinct domains which are predicted to have  
 CC globular regions, cysteine-rich rod-like regions and helical structures.  
 CC The merosin protein has an apparent mol. wt. of about 800 kD and is  
 CC composed of four polypeptides with molecular weights of 300, 200, 200  
 CC and 80 kD. The 300 kD polypeptide is joined to the 200 kD polypeptides  
 CC by disulphide bonds and the 300 and 80 kD polypeptides comprise the  
 CC major subunit protein given in AAR71729. Merosin is found in placenta,  
 CC striated muscle, peripheral nerve, trophoblasts and human Schwann cell  
 CC neoplasms, among other tissues. The 380-400 major merosin subunit also  
 CC yields a 65 kD subunit. The 380-400 merosin subunit has been designated  
 CC merosin polypeptide, merosin subunit, M chain or laminin M chain. These  
 CC peptides may be used as antigens to raise anti-merosin antibodies.  
 CC These antibodies may be used in the detection of merosin, as the absence  
 CC of merosin in a tumour sample indicates malignancy. Contacting a  
 CC neurone with merosin promotes neurite outgrowth. The merosin polypeptide  
 CC may also be used in contacting inhibitors of neurite outgrowth, thereby  
 CC also promoting the outgrowth. Merosin also promotes cell attachment.  
 CC The merosin gene has been mapped to chromosome 6, more precisely to  
 CC bands 6q22->q23.  
 XX  
 SQ Sequence 14 AA;



XX EP52028-A.  
PN  
XX 19-MAY-1982.  
PD  
XX  
PF 14-OCT-1981; 81EP-0401600.  
XX  
XX 14-OCT-1980; 80PR-0021919.  
PR  
XX  
XX (ROUS ) ROUSSEL UCLAF.  
PA (ROUS ) ROUSSEL-UCLAF.  
XX  
XX Gasc JC, Geoffre S, Hospital M, Laurent J;  
PI  
XX  
XX WPI; 1982-41845E/21 (41845E).  
DR  
XX  
PT Hexa:peptide(s) used to treat asthenia, cerebral vascular  
PT troubles etc - Include Cys-Ala-Ala-Phe-D-Lys-Cys,  
PT Cys-Ala-Ala-Cys-D-Lys-Phe and Cys-Glu-His-Cys-D-Lys-Phe, their  
PT di:sulphide(s) and metal complexes  
XX  
XX  
PS Claim 6; Page 16; 20pp; French.  
XX  
XX This peptide is specifically claimed and has the following  
CC properties: It retards loss of learned responses, it retards the  
CC conditioned avoidance response and aids attentiveness and  
CC memory. The peptide can be used in the treatment of memory failure,  
CC senility, cerebral vascular disturbances and mental or nervous  
CC asthenia. See also AAP20155-7 and AAP20407.  
CC  
XX  
XX Sequence 6 AA:  
SQ  
Query Match 29.2%; Score 28; DB 3; Length 6;  
Best Local Similarity 80.0%; Pred. NO. 4.3e+05; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 EHCNF 6  
Db 2 ehckf 6  
RESULT 18  
ID AAY20456 standard; Protein: 9 AA.  
AAZ0456  
XX  
AC AAY20456;  
XX  
XX 22-JUL-1999 (first entry)  
DT  
XX  
XX Human microtubule associated protein 2 mutant fragment 152.  
DE  
XX  
XX Human: beta-amyloid precursor protein; beta-APP; diagnosis; cancer;  
KW frameshift mutation; age-related disease; neurodegenerative disorder;  
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;  
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
KW glial fibrillary acidic protein; GFAP; p53; semaphorin IIR; HUPF-1;  
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
KW high mobility group protein-C; neuroendocrine specific protein A.  
XX  
XX  
XX Synthetic.  
OS  
XX Homo sapiens.  
OS  
XX  
PN WO9845322-A2.  
XX  
XX 15-OCT-1998.  
PD  
XX  
XX 02-APR-1998; 98WO-IB00705.  
PF  
XX  
XX 10-APR-1997; 97US-0043163.  
PR

XX (UYUT-) RIJKSUNIV UTRECHT.  
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
XX  
XX  
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;  
PI  
XX  
XX WPI; 1998-609901/51.  
DR  
XX  
XX N-PSDB; AAX75757.  
DR  
XX  
PT Diagnosing disease by detecting frameshift mutations in RNA or  
PT corresponding protein mutations - used to diagnose cancer and  
PT neurological diseases, particularly Alzheimer's disease, and also  
PT for treatment and prevention with specific ribozymes or wild-type  
PT RNA  
XX  
XX  
PS Disclosure; Figure 6; 258pp; English.  
XX  
XX  
XX This invention describes a novel method for the diagnosis of a disease  
CC caused by, or associated with, an RNA molecule that has a frameshift  
CC mutation. The method is used to diagnose age-related diseases, especially  
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II  
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
CC and many others listed) or susceptibility to these disorders. The method  
CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
CC at an early stage. It is based on the observation that disease may be  
CC caused by mutations in RNA rather than DNA. The invention describes the  
CC used of neuronal system RNA molecules, specifically proteins including  
CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic  
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
CC 2 (bcl-2) proto-oncogene, semaphorin IIR, HUPF-1, high mobility group  
CC protein-C (HMGP-C) and neuroendocrine specific protein A.  
CC  
XX  
XX Sequence 9 AA:  
SQ  
Query Match 29.2%; Score 28; DB 19; Length 9;  
Best Local Similarity 57.1%; Pred. NO. 4.3e+05; Indels 0; Gaps 0;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 FEHCNF 7  
Db 1 fhcsan 7  
RESULT 19  
ID AAM43078 standard; Peptide: 10 AA.  
AAM43078  
XX  
AC AAM43078;  
XX  
XX 22-OCT-2001 (first entry)  
DT  
XX  
XX Mycoplasma genitalium intermolecular complementary peptide, SEQ ID 387.  
DE  
XX  
XX Mycoplasma genitalium; complementary peptide; ligand;  
KW protein-protein interaction; drug design; intermolecular;  
KW intramolecular.  
KW  
XX  
XX  
XX Mycoplasma genitalium.  
OS  
XX  
XX WO200142278-A2.  
PN  
XX  
XX 14-JUN-2001.  
PD  
XX  
XX 13-DEC-2000; 2000WO-GB04778.  
PF  
XX  
XX 13-DEC-1999; 99GB-0029466.  
PR

XX (PROT-) PROTEOM LTD.  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI: 2001-514238/56.  
 XX  
 PT Complementary peptide ligands as reagents and drugs for drug discovery  
 PT programs and as lead ligands to facilitate drug design and development,  
 PT are generated from microbial genome sequences - design and development,  
 XX  
 PS Example 2; Page 103; 161pp; English.  
 XX  
 CC The present sequence is one of a large number of complementary peptide  
 CC ligands generated from Mycoplasma genitalium genome sequences. These  
 CC specific complementary peptides interact with their relevant target  
 CC proteins encoded by the microbial genome. They are capable of  
 CC antagonising or agonising specific interaction of a protein with  
 CC another protein or receptor and are thus useful as reagents and drugs,  
 CC and as lead ligands to facilitate drug design and development. They  
 CC are useful as tools for functional genomic studies, reagents for the  
 CC configuration of high-throughput screens, as a starting point for  
 CC medicinal chemistry manipulation, for peptide mimetics and as  
 CC therapeutic agents. The analysis and acquisition of peptide sequences  
 CC facilitates understanding of protein-protein interactions. The method  
 CC allows for analysis of an entire database at a time, thus overcoming  
 CC sampling problems. The set of complementary peptides includes both  
 CC intermolecular (between proteins) and intramolecular (within a  
 CC protein) sequences.  
 CC  
 SQ Sequence 10 AA;  
 XX  
 Query Match 29.2%; Score 28; DB 22; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 OY 6 FNDVTRLR 14  
 Db 1 fndtlivr 9  
 XX  
 RESULT 20  
 AAY57800  
 ID AAY57800 standard; peptide; 13 AA.  
 XX  
 AC AAY57800;  
 XX  
 DT 20-MAR-2000 (first entry)  
 XX  
 DE TRAM-interaction motif E2P.  
 XX  
 KW Transcriptional adaptor motif; TRAM; TRAM interaction motif; TRIM;  
 KW creb binding protein; transcriptional regulation; cytosolic; antiviral;  
 KW cell cycle inhibitor; viral transcription inhibitor; cancer; tumour;  
 KW viral disease; viral infection; cell cycle; apoptosis; growth arrest.  
 OS Unidentified.  
 XX  
 PN WO9961608-A2.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 26-MAY-1999; 99WO-GB01668.  
 XX  
 PR 26-MAY-1998; 98GB-0011303.  
 PR 05-JAN-1999; 99GB-0000157.  
 XX  
 PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.  
 XX  
 PI O'Connor MJ, Zimmermann H;  
 XX  
 DR WPI: 2000-072620/06.

XX Novel polypeptides cells useful for treating viral disease and cancer  
 PT  
 XX  
 PS Example 1; Fig 2; 73pp; English.  
 XX  
 CC The present invention describes a polypeptide comprising a  
 CC transcriptional adaptor motif (TRAM) or a TRAM-interaction motif  
 CC (TRIM) (I). (I) can be used for identifying compounds (II) comprising a  
 CC polypeptide capable of disrupting an interaction between a TRAM sequence  
 CC and/or a TRIM sequence. Polypeptides which bind to a TRAM sequence or a  
 CC TRIM sequence are identified by incubating, the polypeptide with (I) and  
 CC determining if the polypeptide interacts with (I). Compound (II) is  
 CC useful for preparing a pharmaceutical composition and for disrupting an  
 CC interaction between TRAM sequence and TRIM sequence in vitro, thereby  
 CC inhibiting viral transcription or cell cycle progression in mammalian  
 CC cells especially cancer cell. Compounds which disrupt interaction  
 CC between TRAM/TRAM containing polypeptides can be used therapeutically to  
 CC prevent or treat viral diseases and tumours. The polypeptides reduce  
 CC susceptibility of cells to viral infection and regulate cell cycle  
 CC including apoptosis and growth arrest and can be used to produce  
 CC antibodies against the TRIM or TRAM sequences. HPV types associated  
 CC with high risk or low risk of cervical cancer can be distinguished  
 CC based on the ability of B6 polypeptides to bind to creb binding protein  
 CC (CBP) TRAM sequence. The present sequence represents a TRIM from an  
 CC example from the present invention.  
 CC  
 SQ Sequence 13 AA;  
 XX  
 Query Match 29.2%; Score 28; DB 21; Length 13;  
 Best Local Similarity 57.1%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 OY 4 CNEFDVVT 10  
 Db 3 cdfgdlr 9  
 XX  
 RESULT 21  
 AAR11465  
 ID AAR11465 standard; peptide; 15 AA.  
 XX  
 AC AAR11465;  
 XX  
 DT 11-JUN-1991 (first entry)  
 XX  
 DE Lepidoptera allostatin polypeptide analogue.  
 XX  
 KW LAS; juvenile hormone; insecticide.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1.1  
 FT /label= OTHER  
 FT /note= "pyroglutamine"  
 XX  
 PN EP421935-A.  
 XX  
 PD 10-APR-1991.  
 XX  
 PF 28-SEP-1990; 90EP-0810745.  
 XX  
 PR 02-OCT-1989; 89US-0416097.  
 XX  
 PA (SANO) SANDOZ LTD.  
 XX  
 PI Kramer SJ, Miller CA, Schooley DA, Toschi AG;  
 XX  
 DR WPI: 1991-104135/15.  
 XX  
 PT Polypeptide exhibiting lepidoptera allostatin like activity - and  
 PT inhibiting juvenile hormone synthesis, used as insecticide and  
 PT insect population control agent

PS Claim 6; page 20; 21pp; English.  
 XX This lepidoptera allostatin (LAS) polypeptide analogue is an  
 CC example of a highly generic polypeptide sequence. It exhibits  
 CC LAS-like activity and inhibits juvenile hormone (JH)-I, -II and  
 CC -III activity in lepidoptera. It is used as an insecticide or  
 CC insect population control agent.  
 XX  
 SQ Sequence 15 AA;

Query Match 29.2%; Score 28; DB 12; Length 15;  
 Best Local Similarity 40.0%; Pred. No. 4.8e+02;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 EHCNFDVY 10  
 | | | | |  
 Db 4 frqcytnpis 13

RESULT 22  
 AAW32738  
 ID AAW32738 standard; peptide; 10 AA.  
 XX  
 AC AAW32738;  
 XX  
 DT 30-JAN-1998 (first entry)  
 XX  
 DE Human platelet glycoprotein Ib/IX complex epitope mimotope.  
 XX  
 KW Human platelet glycoprotein Ib/IX complex epitope; probe;  
 KW monoclonal antibody Sz-2; anti mimotope; binding site; antibody;  
 KW platelet adhesion; aggregation; agglutination; von Willebrand factor;  
 KW anti thrombotic drug; ristocetin.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9718236-A1.  
 XX  
 PD 22-MAY-1997.  
 XX  
 PE 08-NOV-1996; 96WO-US17882.  
 XX  
 PR 13-NOV-1995; 95US-0556597.  
 XX  
 PA (UYNV ) UNIV NEW YORK STATE RES FOUND.  
 XX  
 PI Lyle VA, Miller JL;  
 PI  
 PS WPI: 1997-289227/26.  
 DR  
 XX  
 PT Peptide which mimics human platelet glycoprotein Ib/IX complex  
 PT epitope - also anti-mimotope molecule capable of binding to the  
 PT peptide and useful to modulate platelet adhesion, aggregation or  
 PT agglutination  
 XX  
 PS Claim 8; page 99; 115pp; English.  
 XX  
 CC The present sequence represents a mimotope which functionally mimics a  
 CC binding site for a monoclonal antibody (Mab) which recognises an epitope  
 CC within the human platelet glycoprotein (gp) Ib/IX complex. More  
 CC specifically the Mab is Sz-2. The mimotope peptide can be used for  
 CC raising antibodies, as probes to search for anti-mimotopes and to  
 CC neutralise the inhibitory activity of those antibodies which recognise  
 CC the epitope which is mimicked by the peptide. Anti-mimotopes can be used  
 CC to modulate the adhesion, aggregation or agglutination of platelets by  
 CC affecting von Willebrand factor interaction with the platelets through  
 CC the gp Ib/IX receptor, e.g. as an anti-thrombotic drug which inhibits  
 CC the ristocetin induced aggregation of platelets.  
 CC  
 XX Sequence 10 AA;

Query Match 28.6%; Score 27.5; DB 18; Length 10;  
 Best Local Similarity 46.2%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

OY 2 EHCNFDVYTRLR 14  
 : : : : :  
 Db 1 khcs---ihtrlr 10

RESULT 23  
 AAW93135  
 ID AAW93135 standard; peptide; 12 AA.  
 XX  
 AC AAW93135;  
 XX  
 DT 20-MAY-1999 (first entry)  
 XX  
 DE Human growth hormone releasing factor peptide #12.  
 XX  
 KW Growth hormone releasing factor; GRF; human; treatment; osteoporosis;  
 KW hypothalamic pituitary dwarfism; burn; renal failure; infection; wound;  
 KW non-union bone fracture; lactic acidosis; female infertility; cachexia;  
 KW cancer; anabolic; catabolic; T-cell immunodeficiency; tumour; marker;  
 KW neurodegenerative condition; post-surgery.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5854216-A.  
 XX  
 PD 29-DEC-1998.  
 XX  
 PE 24-JUL-1996; 96US-0685357.  
 XX  
 PR 24-JUL-1996; 96US-0685357.  
 PR 23-SEP-1994; 94US-0312244.  
 XX  
 PA (UYMO-) UNIV MONTREAL.  
 XX  
 PI Gaudreau P;  
 PI  
 PS WPI: 1999-094967/08.  
 DR  
 XX  
 PT New polypeptide derivatives are growth hormone releasing factor  
 PT receptor agonists - useful for treatment of e.g. dwarfism, burns,  
 PT osteoporosis, renal failure, infections, infertility, cachexia and  
 PT tumours  
 XX  
 PS Disclosure; Column 39-40; 26pp; English.  
 PS  
 XX This invention describes novel polypeptides of formula RaxRb where  
 CC X = C(O), HNC(=S), HNC(=O)CH2 or CH2; Ra = fluorescein, rhodamine,  
 CC Texas red, a 'body' (RTM), 'Cascade Blue' (RTM), coumarin,  
 CC phycoerythrin, eosin or roseamine; and Rb = a polypeptide of 29 or  
 CC 30 amino acids selected from 26 sequences given in the specification.  
 CC Such peptides are useful for the treatment of hypothalamic pituitary  
 CC dwarfism, burns, osteoporosis, renal failure, non-union bone fracture,  
 CC acute or chronic debilitating illness or infection, wounds, lactation  
 CC failure, infertility in women, cachexia in cancer patients, anabolic  
 CC and/or catabolic problems, T-cell immunodeficiencies, neurodegenerative  
 CC conditions or growth hormone releasing factor (GRF) receptor-dependent  
 CC tumours and for reduction of the incidence of post-surgical problems  
 CC and as markers for GRF receptors.  
 CC  
 XX Sequence 12 AA;

Query Match 28.1%; Score 27; DB 20; Length 12;  
 Best Local Similarity 50.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 8 DVYTRLRNE 17  
 | | | | |

Db 1 dftqlrdd 10

RESULT 24

AAB23282

ID AAB23282 standard; peptide; 12 AA.

AC AAB23282;

DE 02-FEB-2001 (first entry)

XX Hyaluronic acid-binding phage display library peptide, SEQ ID NO.1.

KM Hyaluronic acid-binding peptide; glycosaminoglycan; autoimmune disease;

KM HA-CD4 mediated immune cell migration inhibitor; inflammatory disease;

KM leukocyte infiltration inhibitor; lupus erythematosus;

KM rheumatoid arthritis; graft-versus host disease; infectious disease;

KM hyaluronic acid associated disease; stroke; myocardial infarction;

KM cancer; brain tumour; oedema; glioblastoma multiforme; HIV;

OS Synthetic.

XX WO200057896-A1.

XX 05-OCT-2000.

XX 22-MAR-2000: 2000WO-US07813.

XX 26-MAR-1999: 99US-0126475.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX Takashima A, Mummert ME, Mohamadzaeh M;

PI WPI; 2000-672524/65.

XX

PT Novel composition useful as inhibitor or enhancer of

PT glycosaminoglycan-mediated processes, for treating or preventing

PT infections, inflammatory diseases and cancers -

PS Claim 4; Page 26; 53pp; English.

XX The invention relates to a novel composition comprising an active

XX ingredient which specifically and selectively binds to a natural

XX glycosaminoglycan to alter its normal interactions. The invention also

XX relates to hyaluronic acid (HA)-binding phage display library peptides

XX (AAB23281-823286) which are able to modulate carbohydrate-mediated

XX interactions, and a method of identifying peptides able to modulate

XX carbohydrate-mediated interactions. The composition and peptides

XX modulate HA-CD4 mediated migration of immune cells and inhibit of

XX leukocyte infiltration. The composition of the invention is useful for

XX inhibiting the normal function of natural glycosaminoglycans, preferably

XX hyaluronic acid, chondroitin sulphates A, B, and C, dermatan sulphate,

XX heparin, keratan sulphate, keratansulphate, chitin, chitosan I or chitosan

XX 2. The HA-binding peptides are useful for inhibiting an immune reaction,

XX autoimmune diseases such as lupus erythematosus, rheumatoid arthritis or

XX graft-versus host disease. The HA-binding peptides are also useful for

XX treating diseases associated with hyaluronic acid, such as stroke and

XX myocardial infarction, and/or as an anti-cancer agents, either alone or

XX in combination with chemotherapeutic agents. They can be used to treat

XX oedema associated with brain tumours, particularly oedema associated

XX with glioblastoma multiforme. The peptides are also useful for treating,

XX ameliorating and/or preventing infectious diseases such as HIV

XX infection. They may additionally be used as a contraceptive as HA

XX mediates binding between the sperm and the oocyte. The peptides affords

XX better inhibition of glycosaminoglycan-mediated activity than larger,

XX less specific glycosaminoglycan inhibitors such as receptor antibodies,

XX E.g., anti-CD4 antibodies, and are more specific than chemical

XX inhibitors of glycosaminoglycans such as tummycin and H7. The peptides

XX are significantly smaller than other inhibitors of

XX glycosaminoglycan-mediated activity such as antibodies or soluble

CC glycosaminoglycan, allowing for improved oral and topical formulations,  
CC and are more cost effective to produce. Sequences AAB23282-823286  
CC represent the hyaluronic acid-binding peptide of the invention.  
XX  
SQ Sequence 12 AA;

Query Match 28.1%; Score 27; DB 21; Length 12;

Best Local Similarity 50.0%; Pred. No. 5.5e+02;

Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 HCNFNDVTR 12

Db 3 hmqfnaltr 12

RESULT 25

AAP81811

ID AAP81811 standard; protein; 13 AA.

XX AAP81811;

XX 16-NOV-1990 (first entry)

DE Sequence of peptide with immunological properties and structure of

DE HIV-2 and SIV.1 glycoproteins.

XX Antigen; immunogen; vaccine; diagnostic; AIDS.

XX Immunodeficiency virus.

XX WO8805440-A.

XX 28-JUL-1988.

XX 15-JAN-1988: 88WO-FR00025.

XX 14-APR-1987: 87FR-0005398.

XX (INSP ) INST PASTEUR (ALIZ/).

XX Allzon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M;

PI Tlollais P, Chakrabarti L, Desrosiers R;

XX WPI; 1988-220290/31.

DR New peptide(s) with immunological properties of HIV-2 envelope protein -

XX have the structure of simian immune deficiency virus proteins,

XX useful in diagnosis and of vaccine components

XX Claim 12; Page 44; 86pp; French.

XX New peptides which have immunological properties in common with those of

XX the peptide skeleton of the envelope protein of HIV-2 and also have a

XX antigenic and immunogenic conjugates confg. the peptides and a kit to

XX detect HIV-2 in biological fluids are new. The peptides are useful for in

XX vitro diagnosis of HIV-2 infection and some of them can be used as

XX components of immunogens and vaccines against HIV. Antibodies raised

XX against them can be used for treatment of AIDS.

XX Sequence 13 AA;

SQ

Query Match 28.1%; Score 27; DB 9; Length 13;

Best Local Similarity 50.0%; Pred. No. 6e+02;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 HCNFNDVTR 10

Db 2 hontsvit 9



Tue Apr 2 09:51:44 2002

us-09-020-393b-3\_copy\_42\_58.closed.rag

Page 13

Search completed: April 2, 2002, 09:24:40  
Job time: 109 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:23:26 ; Search time 90.24 Seconds  
(without alignments)  
52.307 Million cell updates/sec

Title: US-09-020-393B-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTTRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 359122

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 43    | 44.8        | 15     | 22 | US-09-835-752-1   |
| 2          | 38    | 39.6        | 17     | 20 | US-09-607-158-9   |
| 3          | 33    | 34.4        | 15     | 6  | US-08-287-535-6   |
| 4          | 32    | 33.3        | 16     | 1  | PCT-US94-08453-11 |
| 5          | 32    | 33.3        | 16     | 6  | US-08-232-513-11  |
| 6          | 32    | 33.3        | 16     | 12 | US-08-824-238-11  |
| 7          | 32    | 33.3        | 16     | 12 | US-08-824-808-11  |
| 8          | 32    | 33.3        | 16     | 14 | US-09-017-259-11  |
| 9          | 32    | 33.3        | 17     | 1  | PCT-US97-04143-16 |

|    |      |      |    |    |                     |                   |
|----|------|------|----|----|---------------------|-------------------|
| 10 | 32   | 33.3 | 17 | 13 | US-08-928-074-16    | Sequence 16, Appl |
| 11 | 32   | 33.3 | 17 | 13 | US-08-928-074-16    | Sequence 16, Appl |
| 12 | 32   | 33.3 | 17 | 13 | US-08-928-074-16    | Sequence 16, Appl |
| 13 | 32   | 33.3 | 17 | 15 | US-09-172-713-11    | Sequence 11, Appl |
| 14 | 30   | 31.2 | 8  | 15 | US-09-187-859-1594  | Sequence 1594, Ap |
| 15 | 30   | 31.2 | 8  | 15 | US-09-187-859-1594  | Sequence 1594, Ap |
| 16 | 30   | 31.2 | 8  | 22 | US-09-839-542-1594  | Sequence 1594, Ap |
| 17 | 30   | 31.2 | 9  | 12 | US-08-821-739-1     | Sequence 1, Appl  |
| 18 | 30   | 31.2 | 9  | 15 | US-09-187-859-1597  | Sequence 1597, Ap |
| 19 | 30   | 31.2 | 9  | 15 | US-09-187-859-1597  | Sequence 1597, Ap |
| 20 | 30   | 31.2 | 9  | 22 | US-09-839-542-1597  | Sequence 1597, Ap |
| 21 | 30   | 31.2 | 16 | 1  | PCT-US96-19065-21   | Sequence 21, Appl |
| 22 | 29   | 30.2 | 13 | 24 | US-60-299-378-28    | Sequence 28, Appl |
| 23 | 29   | 30.2 | 13 | 24 | US-60-300-614-28    | Sequence 28, Appl |
| 24 | 29   | 30.2 | 14 | 15 | US-09-169-404-23    | Sequence 23, Appl |
| 25 | 28   | 29.2 | 9  | 3  | US-07-669-789-4     | Sequence 4, Appl  |
| 26 | 28   | 29.2 | 13 | 13 | US-08-908-371B-4    | Sequence 4, Appl  |
| 27 | 28   | 29.2 | 13 | 21 | US-09-765-101-34    | Sequence 34, Appl |
| 28 | 28   | 29.2 | 10 | 24 | US-60-231-266-34    | Sequence 34, Appl |
| 29 | 27.5 | 28.6 | 10 | 1  | PCT-US96-17882-133  | Sequence 133, App |
| 30 | 27.5 | 28.6 | 10 | 16 | US-09-258-947-133   | Sequence 133, App |
| 31 | 27   | 28.1 | 12 | 8  | US-08-433-108-20    | Sequence 20, Appl |
| 32 | 27   | 28.1 | 12 | 9  | US-08-513-366-20    | Sequence 20, Appl |
| 33 | 27   | 28.1 | 12 | 9  | US-08-513-366A-20   | Sequence 20, Appl |
| 34 | 27   | 28.1 | 12 | 10 | US-08-685-357-12    | Sequence 12, Appl |
| 35 | 27   | 28.1 | 12 | 18 | US-09-444-251-20    | Sequence 20, Appl |
| 36 | 27   | 28.1 | 12 | 19 | US-09-532-709B-1    | Sequence 1, Appl  |
| 37 | 27   | 28.1 | 12 | 20 | US-09-685-010-72    | Sequence 72, Appl |
| 38 | 27   | 28.1 | 14 | 1  | PCT-US00-30036A-112 | Sequence 112, App |
| 39 | 27   | 28.1 | 14 | 1  | PCT-US01-11968-729  | Sequence 729, App |
| 40 | 27   | 28.1 | 14 | 14 | US-09-009-843C-58   | Sequence 58, Appl |
| 41 | 27   | 28.1 | 14 | 14 | US-09-009-843D-58   | Sequence 58, Appl |
| 42 | 27   | 28.1 | 14 | 22 | US-09-833-245-729   | Sequence 729, App |
| 43 | 27   | 28.1 | 15 | 17 | US-09-328-913A-15   | Sequence 15, Appl |
| 44 | 27   | 28.1 | 15 | 21 | US-09-773-416-4     | Sequence 4, Appl  |
| 45 | 27   | 28.1 | 15 | 22 | US-09-819-308-18    | Sequence 18, Appl |
| 46 | 27   | 28.1 | 16 | 15 | US-09-161-094-16    | Sequence 16, Appl |
| 47 | 27   | 28.1 | 16 | 15 | US-09-161-094A-16   | Sequence 16, Appl |
| 48 | 27   | 28.1 | 17 | 10 | US-08-622-338-29    | Sequence 29, Appl |
| 49 | 27   | 28.1 | 17 | 10 | US-08-622-338B-29   | Sequence 29, Appl |
| 50 | 27   | 28.1 | 17 | 10 | US-08-622-338B-29   | Sequence 29, Appl |

#### ALIGNMENTS

RESULT 1  
US-09-835-752-1  
; Sequence 1, Application US/09835752  
; GENERAL INFORMATION:  
; APPLICANT: Halperin, Jose  
; TITLE OF INVENTION: Methods, Products and Treatments for Diabetes  
; FILE REFERENCE: H0498/7137(ERC)  
; CURRENT APPLICATION NUMBER: US/09/835,752  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: US 06/203,254  
; PRIOR FILING DATE: 2000-05-08  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO: 1  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-835-752-1

Query Match 44.8%; Score 43; DB 22; Length 15;  
Best Local Similarity 87.5%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FEHCNFND 8  
DB 7 FEHCNFND 14

RESULT 2  
US-09-607-158-9  
; Sequence 9, Application US/09607158  
; GENERAL INFORMATION:  
; APPLICANT: SAKAKAKA, MASAHIRO  
; APPLICANT: TANAKA, JUNYA  
; APPLICANT: SATO, KOJI  
; APPLICANT: MORITA, FUMIO  
; APPLICANT: SADAMOTO, YASUTAKA  
; TITLE OF INVENTION: CYTOPROTECTIVE AGENTS COMPRISING PROSAPOSIN-RELATED  
; FILE REFERENCE: 49960  
; CURRENT APPLICATION NUMBER: US/09/607,158  
; CURRENT FILING DATE: 2000-06-29  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-607-158-9

Query Match 39.6%; Score 38; DB 20; Length 17;  
Best Local Similarity 60.0%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
DB 2 EHCSLNENIT 11

RESULT 3  
US-08-287-535-6  
; Sequence 6, Application US/08287535  
; GENERAL INFORMATION:  
; APPLICANT: Moe, Gregory R.  
; TITLE OF INVENTION: Conformationally-Stabilized Peptides  
; NUMBER OF INVENTION: and Use Thereof As Immunogens  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Connolly and Hutz  
; STREET: 1220 Market Street  
; CITY: Wilmington  
; STATE: Delaware  
; COUNTRY: U.S.A.  
; ZIP: 19899  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM/PC or Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/287,535  
; FILING DATE: Not yet Assigned  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA: None  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 Amino Acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Peptide  
US-08-287-535-6

Query Match 34.4%; Score 33; DB 6; Length 15;  
Best Local Similarity 45.5%; Pred. No. 3e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CNFNVDVTTLR 14

DB 1 CNFNKLSQLQ 11  
RESULT 4  
PCT-US94-08453-11  
; Sequence 11, Application PC/TUS9408453  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/08453  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: O'BRIEN,002CPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; AMTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
PCT-US94-08453-11

Query Match 33.3%; Score 32; DB 1; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
DB 2 EHCSLNENIT 11

RESULT 5  
US-08-232-513-11  
; Sequence 11, Application US/08232513  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA

COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,513  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'BRIEN, 002CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-232-513-11

Query Match 33.3%; Score 32; DB 6; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFDVTT 11  
| | | : | : |  
Db 2 EHCNFDVTT 11

RESULT 6  
US-08-824-238-11  
Sequence 11, Application US/08824238  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
TITLE OF INVENTION: as Therapeutic Agents  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,238  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: O'BRIEN, 002CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-824-238-11

Query Match 33.3%; Score 32; DB 12; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFDVTT 11  
| | | : | : |  
Db 2 EHCNFDVTT 11

RESULT 7  
US-08-824-808-11  
Sequence 11, Application US/08824808  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
TITLE OF INVENTION: as Therapeutic Agents  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,808  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'BRIEN, 002CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-824-808-11

Query Match 33.3%; Score 32; DB 12; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EHCNFNDVTT 11  
Db 2 EHCNSLNNNT 11

## RESULT 8

US-09-017-259-11  
; Sequence 11, Application US/09017259  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/017,259  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,238  
; FILING DATE:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: O'Brien, 002CPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-09-017-259-11

## Query Match

Best Local Similarity 33.3%; Score 32; DB 14; Length 16;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNFNDVTT 11  
Db 2 EHCNSLNNNT 11

## RESULT 9

PCT-US97-04143-16  
; Sequence 16, Application PC/TUS9704143  
; GENERAL INFORMATION:  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Methods of Alleviating Neuropathic Pain

;; TITLE OF INVENTION: Using Prosaposin-Derived Peptides  
;; NUMBER OF SEQUENCES: 21  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Campbell and Flores  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92122

## COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US97/04143  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: EP-UD 2474  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 17 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
PCT-US97-04143-16

Query Match 33.3%; Score 32; DB 1; Length 17;  
Best Local Similarity 50.0%; Pred. No. 5.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNFNDVTT 11  
Db 2 EHCNSLNNNT 11

## RESULT 10

US-08-928-074-16  
; Sequence 16, Application US/08928074  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,074  
; FILING DATE: 11-SEP-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/611,307  
; FILING DATE: 05-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/04143  
; FILING DATE: 05-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haller, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347

```
REFERENCE/DOCKET NUMBER: 07256/024001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-928-074-16

Query Match
Best Local Similarity 33.3%; Score 32; DB 13; Length 17;
Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 EHCNENDVTT 11
DB 2 EHCSLNEKIT 11

RESULT 11
US-08-928-074-16
Sequence 16, Application US/08928074A
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
FILE REFERENCE: 07256/024001
CURRENT APPLICATION NUMBER: US/08/928,074A
EARLIER FILING DATE: 1997-09-11
EARLIER APPLICATION NUMBER: 08/611,307
EARLIER FILING DATE: 1996-03-05
EARLIER APPLICATION NUMBER: PCT/US97/04143
EARLIER FILING DATE: 1996-03-05
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Artificial
US-08-928-074-16

Query Match
Best Local Similarity 33.3%; Score 32; DB 13; Length 17;
Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 EHCNENDVTT 11
DB 2 EHCSLNEKIT 11

RESULT 12
US-08-928-074-16
Sequence 16, Application US/08928074B
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
FILE REFERENCE: 07256/024001
CURRENT APPLICATION NUMBER: US/08/928,074B
EARLIER FILING DATE: 1997-09-11
EARLIER APPLICATION NUMBER: 08/611,307
EARLIER FILING DATE: 1996-03-05
EARLIER APPLICATION NUMBER: PCT/US97/04143
EARLIER FILING DATE: 1996-03-05
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 17
```

```
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Artificial
US-08-928-074-16

Query Match
Best Local Similarity 33.3%; Score 32; DB 13; Length 17;
Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 EHCNENDVTT 11
DB 2 EHCSLNEKIT 11

RESULT 13
US-09-172-713-11
Sequence 11, Application US/09172713
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/172,713
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,513
FILING DATE: 21-APR-1994
APPLICATION NUMBER: US 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1643
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..17
OTHER INFORMATION: /label= hepo
US-09-172-713-11

Query Match
Best Local Similarity 33.3%; Score 32; DB 15; Length 17;
Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 EHCNENDVTT 11
DB 2 EHCSLNEKIT 11
```

RESULT 14  
US-09-187-859-1594  
; Sequence 1594, Application US/09187859  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187.859  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1594  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
US-09-187-859-1594  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence

Query Match 31.2%; Score 30; DB 15; Length 8;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CINFNDVT 10  
DB 1 CQINDVT 7

RESULT 15  
US-09-187-859-1594  
; Sequence 1594, Application US/09187859A  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187.859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1594  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
US-09-187-859-1594  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence

Query Match 31.2%; Score 30; DB 15; Length 8;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CINFNDVT 10  
DB 1 CQINDVT 7

RESULT 16  
US-09-839-542-1594  
; Sequence 1594, Application US/09839542  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1594  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
US-09-839-542-1594  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence

Query Match 31.2%; Score 30; DB 22; Length 8;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CINFNDVT 10  
DB 1 CQINDVT 7

RESULT 17  
US-08-821-739-1  
; Sequence 1, Application US/08821739  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Cells, Esteban  
; TITLE OF INVENTION: HLA Binding Peptides and Their  
; NUMBER OF SEQUENCES: 149  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/821.739  
; FILING DATE: 20-MAR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,833  
; FILING DATE: 21-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauer  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-0005081US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-821-739-1



Query Match 31.2% Score 30; DB 12; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 HCNFNDVT 10  
| | | | |  
Db 1 HSNINDVT 8

RESULT 18  
US-09-187-859-1597

; Sequence 1597, Application US/09187859  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1597  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence  
US-09-187-859-1597

Query Match 31.2% Score 30; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CNFNDVT 10  
| | | | |  
Db 1 CQINDVT 7

RESULT 19  
US-09-187-859-1597  
; Sequence 1597, Application US/09187859A  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1597  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence  
US-09-187-859-1597

Query Match 31.2% Score 30; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CNFNDVT 10  
| | | | |  
Db 1 CQINDVT 7

RESULT 20  
US-09-839-542-1597  
; Sequence 1597, Application US/09839542  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1597  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence  
US-09-839-542-1597

Query Match 31.2% Score 30; DB 22; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CNFNDVT 10  
| | | | |  
Db 1 CQINDVT 7

RESULT 21  
PCT-US96-19065-21  
; Sequence 21, Application PC/TUS9619065  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Metal-Regulated Transporters and Uses  
; NUMBER OF SEQUENCES: 21  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/19065  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/018,578  
; FILING DATE: 29-MAY-1996  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silveri, Jean M.  
; REGISTRATION NUMBER: 39,030  
; REFERENCE/DOCKET NUMBER: DCI-099CPC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: Internal

PCT-US96-19065-21

## Query Match

Best Local Similarity 31.2%; Score 30; DB 1; Length 16;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;Qy 7 NDVYTRLENE 17  
| | | | : | : |  
Db 3 NDVYTRLENE 13

## RESULT 22

US-60-299-378-28

; Sequence 28, Application US/60299378

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOGY T

; TITLE OF INVENTION: Ether-a-90-90 FAMILY, HEAG2, EXPRESSED SPECIFICALLY IN THALAMUS,

; FILE REFERENCE: D0123

; CURRENT APPLICATION NUMBER: US/60/299,378

; CURRENT FILING DATE: 2001-06-19

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 28

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-299-378-28

## Query Match

Best Local Similarity 30.2%; Score 29; DB 24; Length 13;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;Qy 4 CNFNDVT 10  
| | | | : | : |  
Db 5 CTFKDT 11

## RESULT 23

US-60-300-614-28

; Sequence 28, Application US/60300614

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOGY T

; TITLE OF INVENTION: Ether-a-90-90 FAMILY, HEAG2, EXPRESSED SPECIFICALLY IN THALAMUS,

; FILE REFERENCE: D0123 PSPI

; CURRENT APPLICATION NUMBER: US/60/300,614

; CURRENT FILING DATE: 2001-06-25

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 28

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-300-614-28

## Query Match

Best Local Similarity 30.2%; Score 29; DB 24; Length 13;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;Qy 4 CNFNDVT 10  
| | | | : | : |  
Db 5 CTFKDT 11

## RESULT 24

US-09-169-404-23

; Sequence 23, Application US/09169404

## GENERAL INFORMATION:

APPLICANT: Engvall, Eva

APPLICANT: Lelivo, Ilmo

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

TITLE OF INVENTION: Fragments and Uses Thereof

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESS: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/169,404

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/125,077

FILING DATE: 22-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US 94/10730

FILING DATE: 21-SEP-1994

APPLICATION DATA:

APPLICATION NUMBER: US 07/472,319

FILING DATE: 30-JAN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/919,951

FILING DATE: 27-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LA 9721

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 23:

LENGTH: 14 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-09-169-404-23

## Query Match

Best Local Similarity 30.2%; Score 29; DB 15; Length 14;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;Qy 4 CNFNDVT 16  
| | | | : | : |  
Db 1 CSTVDITNEN 13

## RESULT 25

US-07-669-789-4

; Sequence 4, Application US/07669789

; GENERAL INFORMATION:

; APPLICANT: Liu, Chung-Cheng

; TITLE OF INVENTION: Use of Purified Ubiquitin Hydrolase in Recovering Polypepti

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

ADDRESS: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/669,789  
FILING DATE: 19910315  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/210,909  
FILING DATE: June 24, 1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 513D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/266-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-669-789-4

Query Match 29.2%; Score 28; DB 3; Length 9;  
Best Local Similarity 71.4%; Pred. No. 2.9e+06; Indels 0; Gaps 0;  
Matches 5; Conservative 2; Mismatches 0;  
QY 6 FNDVTR 12  
|:|:|:|:  
Db 3 FDDVTTQ 9

Search completed: April 2, 2002, 09:26:37  
Job time: 191 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:23:56 ; Search time 15.54 Seconds  
(without alignments)  
78.189 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFDVTRIRENE 17

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 348507 seqs, 71473746 residues  
Total number of hits satisfying chosen parameters: 93745

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 31    | 32.3        | 15     | 6     | US-09-989-919-123   |
| 2          | 30    | 31.2        | 8      | 7     | US-10-006-869-1594  |
| 3          | 30    | 31.2        | 9      | 7     | US-10-006-869-1597  |
| 4          | 30    | 31.2        | 16     | 6     | US-09-579-174-21    |
| 5          | 29    | 30.2        | 11     | 6     | US-09-754-831A-22   |
| 6          | 29    | 29.2        | 10     | 6     | US-09-573-822C-387  |
| 7          | 28    | 29.2        | 13     | 1     | PCT-US01-27692A-5   |
| 8          | 28    | 29.2        | 13     | 6     | US-09-701-080A-5    |
| 9          | 28    | 29.2        | 13     | 6     | US-09-701-080A-10   |
| 10         | 27    | 28.1        | 12     | 6     | US-09-965-738-260   |
| 11         | 27    | 28.1        | 14     | 1     | PCT-US02-08123-1760 |
| 12         | 27    | 28.1        | 16     | 7     | US-10-038-612-16    |
| 13         | 27    | 28.1        | 17     | 7     | US-10-038-612-114   |
| 14         | 26    | 27.1        | 12     | 4     | US-08-816-454B-208  |
| 15         | 26    | 27.1        | 12     | 6     | US-09-921-822-33    |
| 16         | 26    | 27.1        | 12     | 6     | US-09-854-385-373   |
| 17         | 26    | 27.1        | 12     | 6     | US-10-084-813-422   |
| 18         | 26    | 27.1        | 15     | 6     | US-09-828-574-7     |
| 19         | 26    | 27.1        | 15     | 7     | US-10-084-813-419   |
| 20         | 25    | 26.0        | 7      | 7     | US-10-006-869-1592  |
| 21         | 25    | 26.0        | 8      | 1     | PCT-US01-27702A-63  |
| 22         | 25    | 26.0        | 8      | 5     | US-09-918-873-15    |
| 23         | 25    | 26.0        | 10     | 6     | US-09-572-270B-50   |
| 24         | 25    | 26.0        | 10     | 6     | US-09-572-270B-54   |
| 25         | 26    | 0           | 10     | 6     | US-09-572-270B-58   |

|    |    |      |    |   |                     |                    |
|----|----|------|----|---|---------------------|--------------------|
| 26 | 25 | 26.0 | 10 | 6 | US-09-572-270B-62   | Sequence 62, Appl  |
| 27 | 25 | 26.0 | 10 | 6 | US-09-572-270B-66   | Sequence 66, Appl  |
| 28 | 25 | 26.0 | 10 | 6 | US-09-572-270B-70   | Sequence 70, Appl  |
| 29 | 25 | 26.0 | 10 | 6 | US-09-832-230-67    | Sequence 67, Appl  |
| 30 | 25 | 26.0 | 10 | 6 | US-09-572-404B-549  | Sequence 549, Appl |
| 31 | 25 | 26.0 | 10 | 6 | US-09-572-404B-550  | Sequence 550, Appl |
| 32 | 25 | 26.0 | 10 | 6 | US-09-572-404B-3662 | Sequence 3662, Ap  |
| 33 | 25 | 26.0 | 10 | 6 | US-09-572-404B-3663 | Sequence 3663, Ap  |
| 34 | 25 | 26.0 | 10 | 6 | US-09-832-230A-67   | Sequence 67, Appl  |
| 35 | 25 | 26.0 | 10 | 6 | US-10-083-768-67    | Sequence 67, Appl  |
| 36 | 25 | 26.0 | 13 | 6 | US-09-952-677-8     | Sequence 8, Appl   |
| 37 | 25 | 26.0 | 14 | 7 | US-10-005-549-28    | Sequence 28, Appl  |
| 38 | 25 | 26.0 | 15 | 6 | US-09-443-986A-1    | Sequence 1, Appl   |
| 39 | 25 | 26.0 | 15 | 6 | US-09-711-161-14    | Sequence 14, Appl  |
| 40 | 25 | 26.0 | 17 | 6 | US-09-834-784-848   | Sequence 848, Appl |
| 41 | 25 | 26.0 | 17 | 6 | US-09-350-641C-848  | Sequence 848, Appl |
| 42 | 24 | 25.0 | 9  | 7 | US-10-084-813-423   | Sequence 423, Appl |
| 43 | 24 | 25.0 | 11 | 7 | US-10-020-269-26    | Sequence 26, Appl  |
| 44 | 24 | 25.0 | 12 | 7 | US-10-084-813-420   | Sequence 420, Appl |
| 45 | 24 | 25.0 | 12 | 7 | US-10-084-813-421   | Sequence 421, Appl |
| 46 | 24 | 25.0 | 13 | 6 | US-09-419-901A-4    | Sequence 4, Appl   |
| 47 | 24 | 25.0 | 15 | 7 | US-10-038-407-13    | Sequence 13, Appl  |
| 48 | 24 | 25.0 | 15 | 7 | US-10-084-813-417   | Sequence 417, Appl |
| 49 | 24 | 25.0 | 15 | 7 | US-10-084-813-418   | Sequence 418, Appl |
| 50 | 24 | 25.0 | 17 | 6 | US-09-350-641C-1523 | Sequence 1523, Ap  |

## ALIGNMENTS

```
RESULT 1
US-09-989-919-123
; Sequence 123, Application US/09989919
; GENERAL INFORMATION:
; APPLICANT: Medina, Roberto
; APPLICANT: Requion, Hevra
; APPLICANT: Pluta, Jason
; APPLICANT: Ghosh, Malavika
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and
; FILE REFERENCE: DEX-0289
; CURRENT APPLICATION NUMBER: US/09/989,919
; PRIOR APPLICATION NUMBER: 60/252,505
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patent version 3.1
; SEQ ID NO 123
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-919-123

Query Match          32.3%; Score 31; DB 6; Length 15;
Best Local Similarity 30.0%; Pred. No. 56;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 FEHCNFDVTR 10
Db 4 YKHCHMNNLS 13

RESULT 2
US-10-006-869-1594
; Sequence 1594, Application US/10006869
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; CADHERIN-MEDIATED FUNCTIONS
```

FILE REFERENCE: 100086.407C7  
CURRENT APPLICATION NUMBER: US/10/006,869  
CURRENT FILING DATE: 2001-12-03  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1594  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Representative cyclic modulating agent based on  
US-10-006-869-1594

Query Match  
Best Local Similarity 31.2%; Score 30; DB 7; Length 8;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CNFNVT 10  
1 COINDVT 7

RESULT 3  
US-10-006-869-1597  
Sequence 1597, Application US/10006869  
GENERAL INFORMATION:  
APPLICANT: Blaschuk, Orest W.  
APPLICANT: Symonds, James Matthew  
APPLICANT: Gour, Barbara J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
FILE REFERENCE: 100086.407C7  
CURRENT APPLICATION NUMBER: US/10/006,869  
CURRENT FILING DATE: 2001-12-03  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1597  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Representative cyclic modulating agent based on  
US-10-006-869-1597

Query Match  
Best Local Similarity 31.2%; Score 30; DB 7; Length 9;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CNFNVT 10  
1 COINDVT 7

RESULT 4  
US-09-579-174-21  
Sequence 21, Application US/09579174  
GENERAL INFORMATION:  
APPLICANT: Gueriot, Mary Lou et al.  
TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR  
FILE REFERENCE: DCI-099CDDV  
CURRENT APPLICATION NUMBER: US/09/579,174  
CURRENT FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 09/107,858  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: 08/758,621  
PRIOR FILING DATE: 1996-11-27  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 21

LENGTH: 16  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-579-174-21

Query Match  
Best Local Similarity 31.2%; Score 30; DB 6; Length 16;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTRLRENE 17  
3 NDVTLPIKEDD 13

RESULT 5  
US-09-754-831A-22  
Sequence 22, Application US/09754831A  
GENERAL INFORMATION:  
APPLICANT: Oppermann, Herman  
APPLICANT: Kuberassampath, Thangavel  
APPLICANT: Rueger, David  
TITLE OF INVENTION: Osteogenic Devices  
FILE REFERENCE: STR-008CN  
CURRENT APPLICATION NUMBER: US/09/754,831A  
CURRENT FILING DATE: 2001-01-03  
PRIOR APPLICATION NUMBER: US 08/375,901  
PRIOR FILING DATE: 1995-01-20  
PRIOR APPLICATION NUMBER: US 08/145,812  
PRIOR FILING DATE: 1993-11-01  
PRIOR APPLICATION NUMBER: US 07/995,345  
PRIOR FILING DATE: 1992-12-22  
PRIOR APPLICATION NUMBER: US 07/315,342  
PRIOR FILING DATE: 1989-02-23  
PRIOR APPLICATION NUMBER: US 07/232,630  
PRIOR FILING DATE: 1988-08-15  
PRIOR APPLICATION NUMBER: US 07/179,406  
PRIOR FILING DATE: 1988-04-08  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 22  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: peptide fragment  
US-09-754-831A-22

Query Match  
Best Local Similarity 30.2%; Score 29; DB 6; Length 11;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 PEHCNEND 8  
1 FLHCQFSE 8

RESULT 6  
US-09-573-822C-387  
Sequence 387, Application US/09573822C  
GENERAL INFORMATION:  
APPLICANT: Proteom Ltd  
TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome  
FILE REFERENCE: Microbe patent  
CURRENT APPLICATION NUMBER: US/09/573,822C  
CURRENT FILING DATE: 2000-05-18  
NUMBER OF SEQ ID NOS: 804  
SOFTWARE: ProclPatent version 1.0  
SEQ ID NO 387  
LENGTH: 10  
TYPE: PRT

ORGANISM: mycoplasma genitalium  
FEATURE:  
OTHER INFORMATION: Sequence located in MG121 at 123-132 and may interact with Sequen  
US-09-573-822C-387

Query Match 29.2%; Score 28; DB 6; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 FNDVTRLR 14  
|:|:|:|:  
Db 1 FSDTTLRVR 9

RESULT 7  
PCT-US01-27692A-5  
Sequence 5, Application PC/TUS0127692A  
GENERAL INFORMATION:  
APPLICANT: Board of Regents, The University of Texas System  
TITLE OF INVENTION: Human and Mouse Targeting Peptides Identified by Phage Display  
FILE REFERENCE: 005774.P003PCT  
CURRENT APPLICATION NUMBER: PCT/US01/27692A  
CURRENT FILING DATE: 2001-09-07  
NUMBER OF SEQ ID NOS: 251  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 5  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: (1):(13)  
OTHER INFORMATION: synthetic construct  
PCT-US01-27692A-5

Query Match 29.2%; Score 28; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FEHC 4  
|:|:|:|:  
Db 10 FEHC 13

RESULT 8  
US-09-701-080A-5  
Sequence 5, Application US/09701080A  
GENERAL INFORMATION:  
APPLICANT: O'CONNOR, MARK J.  
APPLICANT: ZIMMERMAN, HOLGER  
TITLE OF INVENTION: POLYPEPTIDES FROM CREB BINDING PROTEIN AND RELATED PROTEIN P300 F  
FILE REFERENCE: 117-328  
CURRENT APPLICATION NUMBER: US/09/701,080A  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: GB 9811303.8  
PRIOR FILING DATE: 1998-05-26  
PRIOR APPLICATION NUMBER: GB 9900157.0  
PRIOR FILING DATE: 1999-01-05  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 5  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:derived from E2F  
US-09-701-080A-5

Query Match 29.2%; Score 28; DB 6; Length 13;  
Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNEFDVTR 10  
|:|:|:|:  
Db 3 CDFGDLTR 9

RESULT 9  
US-09-701-080A-10  
Sequence 10, Application US/09701080A  
GENERAL INFORMATION:  
APPLICANT: O'CONNOR, MARK J.  
APPLICANT: ZIMMERMAN, HOLGER  
TITLE OF INVENTION: POLYPEPTIDES FROM CREB BINDING PROTEIN AND RELATED PROTEIN P30  
FILE REFERENCE: 117-328  
CURRENT APPLICATION NUMBER: US/09/701,080A  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: GB 9811303.8  
PRIOR FILING DATE: 1998-05-26  
PRIOR APPLICATION NUMBER: GB 9900157.0  
PRIOR FILING DATE: 1999-01-05  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 10  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:derived from E2F  
US-09-701-080A-10

Query Match 29.2%; Score 28; DB 6; Length 13;  
Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CNEFDVTR 10  
|:|:|:|:  
Db 3 CDFGDLTR 9

RESULT 10  
US-09-965-738-260  
Sequence 260, Application US/09965738  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy  
TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnosti  
FILE REFERENCE: 40715-258841  
CURRENT APPLICATION NUMBER: US/09/965,738  
CURRENT FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US 60/284,175  
PRIOR FILING DATE: 2001-04-17  
NUMBER OF SEQ ID NOS: 306  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 260  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-965-738-260

Query Match 28.1%; Score 27; DB 6; Length 12;  
Best Local Similarity 45.5%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 FEHCNFDVTR 11  
|:|:|:|:  
Db 1 FTRHNFVPTTS 11

RESULT 11  
PCT-US02-08123-1760  
; Sequence 1760, Application PC/TUS0208123  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS904PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/08123  
; CURRENT FILING DATE: 2002-03-19  
; PRIOR APPLICATION NUMBER: US 60/331,287  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/306,171  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: US 60/277,340  
; PRIOR FILING DATE: 2001-03-21  
; NUMBER OF SEQ ID NOS: 2048  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1760  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-08123-1760

Query Match 28.1%; Score 27; DB 1; Length 14;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 NDVTRL 13  
Db 5 NDVTKL 11

RESULT 12  
US-10-038-612-16  
; Sequence 16, Application US/10038612  
; GENERAL INFORMATION:  
; APPLICANT: Ben-Sasson, Shmuel A.  
; TITLE OF INVENTION: Short Peptides Which Selectively  
; FILE REFERENCE: 1242.1029-000 (CMCC-679)  
; CURRENT APPLICATION NUMBER: US/10/038,612  
; CURRENT FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 09/161,094  
; PRIOR FILING DATE: 1998-09-25  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: unknown  
; FEATURE:  
; OTHER INFORMATION: CK IIA  
US-10-038-612-16

Query Match 28.1%; Score 27; DB 7; Length 16;  
Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 FEHCNFD 8  
Db 1 FEHVNTD 8

RESULT 13  
US-10-038-612-114  
; Sequence 114, Application US/10038612  
; GENERAL INFORMATION:  
; APPLICANT: Ben-Sasson, Shmuel A.  
; TITLE OF INVENTION: Short Peptides Which Selectively  
; TITLE OF INVENTION: Modulate the Activity of Protein Kinases

; FILE REFERENCE: 1242.1029-000 (CMCC-679)  
; CURRENT APPLICATION NUMBER: US/10/038,612  
; CURRENT FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 09/161,094  
; PRIOR FILING DATE: 1998-09-25  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 114  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MYRISTATE  
; LOCATION: (1)-(10)  
; NAME/KEY: AMIDATION  
; LOCATION: (0)...(17)  
; OTHER INFORMATION: CK IIA  
US-10-038-612-114

Query Match 28.1%; Score 27; DB 7; Length 17;  
Best Local Similarity 62.5%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 FEHCNFD 8  
Db 2 FEHVNTD 9

RESULT 14  
US-08-816-454B-208  
; Sequence 208, Application US/08816454B  
; GENERAL INFORMATION:  
; APPLICANT: SLINGLUFF, Craig L.  
; HUNT, Donald F.  
; SHABANOWITZ, Jeffrey  
; COX, Andrea L.  
; ENGELHARD, Victor H.  
; KITTELSEN, David  
; SKIPPER, Jonathan  
; HENDRICKSON, Ronald C.  
; TITLE OF INVENTION: PEPTIDES RECOGNIZED BY MELANOMA-SPECIFIC  
; A1-, A2, A3-RESTRICTED CYTOTOXIC LYMPHOCYTES AND USES  
; THEREFOR  
; NUMBER OF SEQUENCES: 294  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NIEMARK, P.L.L.C.  
; STREET: 624 Ninth Street N.W., Ste. 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/816,454B  
; FILING DATE: 12-Mar-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/027,627  
; FILING DATE: 04-OCT-1996  
; APPLICATION NUMBER: US 60/013,972  
; FILING DATE: 19-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOPER, Iver P.  
; REGISTRATION NUMBER: 28,005  
; REFERENCE/DOCKET NUMBER: SLINGLUFF-2/3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-5197  
; TELEFAX: (202) 737-3528



9  
; INFORMATION FOR SEQ ID NO: 208:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 208:  
US-08-816-454B-208

Query Match 27.1%; Score 26; DB 4; Length 12;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 FNDVYT 11  
111:1  
DB 6 FNDINT 11

## RESULT 15

US-09-921-822-33  
; Sequence 33, Application US/09921822

; GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.

Miller, Jeff F.

Cull, Millard G.

Stemmer, Willem P.C.

Gates, Christian M.

TITLE OF INVENTION: Peptide Library and Screening Method

NUMBER OF SEQUENCES: 162

CORRESPONDENCE ADDRESSES:

ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/921,822

FILING DATE: 02-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/504,787

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/290,641

FILING DATE: 15-Aug-1994

APPLICATION NUMBER: US 07/963,321

FILING DATE: 15-Oct-1992

ATTORNEY/AGENT INFORMATION:

NAME: Smith, William M.

REGISTRATION NUMBER: 30,223

REFERENCE/DOCKET NUMBER: 16528J-001240US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-326-2400

TELEFAX: 415-326-2422

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

CLONE: 7 4 1.1

SEQUENCE DESCRIPTION: SEQ ID NO: 33:

US-09-921-822-33

Query Match 27.1%; Score 26; DB 6; Length 12;  
Best Local Similarity 55.6%; Pred. No. 3.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VTTRLRNE 17  
111111:  
DB 1 VTQVRVSNK 9

## RESULT 16

US-09-954-385-373

; Sequence 373, Application US/09954385

; GENERAL INFORMATION:

APPLICANT: Aehle, Wolfgang

APPLICANT: Baldwin, Toby L.

APPLICANT: Van Gastel, Franciscus J.C.

APPLICANT: Janssen, Giselle G.

APPLICANT: Murray, Christopher J.

APPLICANT: Wang, Huang

APPLICANT: Winetky, Deborah S.

TITLE OF INVENTION: Binding Phenol Oxidizing Enzyme-peptide

TITLE OF INVENTION: Complexes

FILE REFERENCE: GC690

CURRENT APPLICATION NUMBER: US/09/954,385

CURRENT FILING DATE: 2001-09-12

NUMBER OF SEQ ID NOS: 433

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 373

LENGTH: 12

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: binding peptide

US-09-954-385-373

Query Match 27.1%; Score 26; DB 6; Length 12;  
Best Local Similarity 50.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 NFNDVYTR 12  
111111:  
DB 5 NFSNVSTR 12

## RESULT 17

US-10-084-813-422

; Sequence 422, Application US/10084813

; GENERAL INFORMATION:

APPLICANT: SAXINGER, CARL

TITLE OF INVENTION: POLYPEPTIDES THAT BIND HIV GP120 AND RELATED NUCLEIC

FILE REFERENCE: 215875

CURRENT APPLICATION NUMBER: US/10/084,813

CURRENT FILING DATE: 2002-02-27

PRIOR APPLICATION NUMBER: PCT/US00/23505

PRIOR FILING DATE: 2000-08-25

PRIOR APPLICATION NUMBER: US 60/151,270

PRIOR FILING DATE: 1999-08-27

NUMBER OF SEQ ID NOS: 1242

SOFTWARE: Patentin version 3.1

SEQ ID NO 422

LENGTH: 12

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: binding peptide

US-10-084-813-422

Query Match 27.1%; Score 26; DB 7; Length 12;  
Best Local Similarity 50.0%; Pred. No. 3.1e+02;

REFERENCE: A33864 090495.0232

;; CURRENT APPLICATION NUMBER: US/09/918,873  
;; CURRENT FILING DATE: 2002-07-31  
;; NUMBER OF SEQ ID NOS: 39  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 15  
;; LENGTH: 8  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; DATABASE ACCESSION NUMBER: 18158777  
;; DATABASE ENTRY DATE: 2001-01-29  
US-09-918-873-15

Query Match 26.0%; Score 25; DB 5; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.9e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 EHON 5  
DB 2 DHON 5

RESULT 23  
US-09-572-270B-50  
;; Sequence 50, Application US/09572270B  
;; GENERAL INFORMATION:  
;; APPLICANT: Proteom Ltd  
;; TITLE OF INVENTION: Complementary peptide ligands from plant genomes  
;; FILE REFERENCE: Plant patent  
;; CURRENT APPLICATION NUMBER: US/09/572,270B  
;; CURRENT FILING DATE: 2000-05-17  
;; NUMBER OF SEQ ID NOS: 1144  
;; SOFTWARE: ProPatent version 1.0  
;; SEQ ID NO: 50  
;; LENGTH: 10  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis Thaliana  
;; FEATURE:  
;; OTHER INFORMATION: Sequence located in CADI. at 36-45 and may interact with Sequence  
US-09-572-270B-50

Query Match 26.0%; Score 25; DB 6; Length 10;  
Best Local Similarity 57.1%; Pred. No. 3.7e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
DB 3 NDVTVKI 9

RESULT 24  
US-09-572-270B-54  
;; Sequence 54, Application US/09572270B  
;; GENERAL INFORMATION:  
;; APPLICANT: Proteom Ltd  
;; TITLE OF INVENTION: Complementary peptide ligands from plant genomes  
;; FILE REFERENCE: Plant patent  
;; CURRENT APPLICATION NUMBER: US/09/572,270B  
;; CURRENT FILING DATE: 2000-05-17  
;; NUMBER OF SEQ ID NOS: 1144  
;; SOFTWARE: ProPatent version 1.0  
;; SEQ ID NO: 54  
;; LENGTH: 10  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis Thaliana  
;; FEATURE:  
;; OTHER INFORMATION: Sequence located in CADI. at 36-45 and may interact with Sequence  
US-09-572-270B-54

Query Match 26.0%; Score 25; DB 6; Length 10;  
Best Local Similarity 57.1%; Pred. No. 3.7e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
DB 3 NDVTVKI 9

RESULT 25  
US-09-572-270B-58  
;; Sequence 58, Application US/09572270B  
;; GENERAL INFORMATION:  
;; APPLICANT: Proteom Ltd  
;; TITLE OF INVENTION: Complementary peptide ligands from plant genomes  
;; FILE REFERENCE: Plant patent  
;; CURRENT APPLICATION NUMBER: US/09/572,270B  
;; CURRENT FILING DATE: 2000-05-17  
;; NUMBER OF SEQ ID NOS: 1144  
;; SOFTWARE: ProPatent version 1.0  
;; SEQ ID NO: 58  
;; LENGTH: 10  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis Thaliana  
;; FEATURE:  
;; OTHER INFORMATION: Sequence located in CADI. at 36-45 and may interact with Sequence  
US-09-572-270B-58

Query Match 26.0%; Score 25; DB 6; Length 10;  
Best Local Similarity 57.1%; Pred. No. 3.7e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 NDVTTRL 13  
DB 3 NDVTVKI 9

Search completed: April 2, 2002, 09:26:59  
Job time: 183 sec

Tue Apr 2 09:51:50 2002

us-09-020-393b-3\_copy-42\_58\_closed.ram

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2002, 09:22:51 ; Search time 12.51 Seconds  
(without alignments)  
30.580 Million cell updates/sec

Title: US-09-020-393b-3\_COPY\_42\_58  
Perfect score: 96  
Sequence: 1 FEHCNFNDVTRRLRENE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 100354

Minimum DB seq length: 0  
Maximum DB seq length: 17

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 32    | 33.3        | 17     | 1     | US-08-232-513A-11  |
| 2          | 32    | 33.3        | 17     | 4     | US-09-231-159-16   |
| 3          | 32    | 33.3        | 17     | 4     | US-08-611-307-16   |
| 4          | 30    | 31.2        | 16     | 2     | US-08-758-621-21   |
| 5          | 30    | 31.2        | 16     | 4     | US-09-107-858-21   |
| 6          | 29    | 30.2        | 11     | 6     | 5496552-8          |
| 7          | 29    | 30.2        | 14     | 2     | US-08-460-309-23   |
| 8          | 29    | 30.2        | 14     | 2     | US-08-125-077-23   |
| 9          | 29    | 30.2        | 14     | 6     | 5444158-4          |
| 10         | 28    | 29.2        | 15     | 1     | US-08-522-326-6    |
| 11         | 27.5  | 28.6        | 10     | 2     | US-08-556-597-133  |
| 12         | 27    | 28.1        | 12     | 2     | US-08-685-357B-12  |
| 13         | 27    | 28.1        | 12     | 3     | US-08-952-568-20   |
| 14         | 27    | 28.1        | 16     | 6     | 5194585-1          |
| 15         | 27    | 28.1        | 16     | 6     | 5194585-3          |
| 16         | 26.5  | 27.6        | 15     | 1     | US-08-218-025A-105 |
| 17         | 26    | 27.1        | 11     | 4     | US-09-224-785-22   |
| 18         | 26    | 27.1        | 11     | 4     | US-09-224-785-25   |
| 19         | 26    | 27.1        | 12     | 1     | US-07-778-233B-33  |
| 20         | 26    | 27.1        | 12     | 1     | US-07-963-321-33   |
| 21         | 26    | 27.1        | 12     | 1     | US-08-290-641-33   |
| 22         | 26    | 27.1        | 12     | 1     | US-08-548-540-33   |
| 23         | 26    | 27.1        | 12     | 5     | PCT-US96-09809-33  |
| 24         | 26    | 27.1        | 14     | 2     | US-08-413-708B-5   |
| 25         | 26    | 27.1        | 14     | 3     | US-09-192-048-24   |
| 26         | 27    | 27.1        | 16     | 4     | US-09-164-186-5    |
| 27         | 26    | 27.1        | 17     | 4     | US-08-706-344C-15  |

|    |      |      |    |   |                    |                   |
|----|------|------|----|---|--------------------|-------------------|
| 28 | 25.5 | 26.6 | 12 | 2 | US-08-468-812-14   | Sequence 14, Appl |
| 29 | 25.5 | 26.6 | 12 | 4 | US-08-590-563-22   | Sequence 22, Appl |
| 30 | 25   | 26.0 | 6  | 1 | US-08-375-911A-5   | Sequence 5, Appl1 |
| 31 | 25   | 26.0 | 9  | 4 | US-09-502-600-69   | Sequence 69, Appl |
| 32 | 25   | 26.0 | 10 | 2 | US-08-764-640-67   | Sequence 67, Appl |
| 33 | 25   | 26.0 | 10 | 3 | US-08-973-225-67   | Sequence 67, Appl |
| 34 | 25   | 26.0 | 10 | 3 | US-09-244-298A-67  | Sequence 67, Appl |
| 35 | 25   | 26.0 | 10 | 4 | US-09-516-704-67   | Sequence 67, Appl |
| 36 | 25   | 26.0 | 11 | 4 | US-09-224-785-29   | Sequence 29, Appl |
| 37 | 25   | 26.0 | 13 | 4 | US-08-836-567-17   | Sequence 17, Appl |
| 38 | 25   | 26.0 | 13 | 4 | US-09-196-390-8    | Sequence 8, Appl1 |
| 39 | 25   | 26.0 | 15 | 1 | US-08-221-583-17   | Sequence 17, Appl |
| 40 | 25   | 26.0 | 15 | 1 | US-08-221-583-18   | Sequence 18, Appl |
| 41 | 25   | 26.0 | 15 | 1 | US-08-221-583-19   | Sequence 19, Appl |
| 42 | 25   | 26.0 | 15 | 1 | US-08-310-341B-1   | Sequence 1, Appl1 |
| 43 | 25   | 26.0 | 15 | 5 | PCT-US95-04018-17  | Sequence 17, Appl |
| 44 | 25   | 26.0 | 15 | 5 | PCT-US95-04018-18  | Sequence 18, Appl |
| 45 | 25   | 26.0 | 15 | 5 | PCT-US95-05657-1   | Sequence 19, Appl |
| 46 | 25   | 26.0 | 17 | 5 | PCT-US95-05657-1   | Sequence 1, Appl1 |
| 47 | 25   | 26.0 | 17 | 4 | US-09-082-279B-848 | Sequence 848, App |
| 48 | 24   | 25.0 | 8  | 2 | US-08-672-610A-53  | Sequence 53, Appl |
| 49 | 24   | 25.0 | 8  | 5 | PCT-US91-02166-14  | Sequence 14, Appl |
| 50 | 24   | 25.0 | 9  | 4 | US-09-518-046-74   | Sequence 74, Appl |

#### ALIGNMENTS

RESULT 1  
US-08-232-513A-11  
; Sequence 11, Application US/08232513A  
; Patent No. 5700909  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Protoposin and Cytokine-Derived Peptides  
; NUMBER OF INVENTIONS: as Therapeutic Agents  
; CORRESPONDENCE ADDRESSES: 20  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,513A  
; FILING DATE: 21-Apr-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UD 1643  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9901  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..17  
; OTHER INFORMATION: /label= hepo

US-08-232-513A-11

Query Match 33.3%; Score 32; DB 1; Length 17;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11  
Db 2 EHCNENDVTT 11

RESULT 2

US-09-231-159-16  
Sequence 16, Application US/09231159  
Patent No. 6268347

GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Methods of Alleviating Neuropathic Pain  
TITLE OF INVENTION: Using Prosaposin-Derived Peptides  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/231,159  
FILING DATE:

CLASSIFICATION:

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/231,159  
FILING DATE: 05-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-09-231-159-16

Query Match 33.3%; Score 32; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11  
Db 2 EHCNENDVTT 11

RESULT 3

US-08-611-307-16  
Sequence 16, Application US/08611307  
Patent No. 6271196

GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Methods of Alleviating Neuropathic Pain  
TITLE OF INVENTION: Using Prosaposin-Derived Peptides  
NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/611,307

FILING DATE: 05-MAR-1996

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-UD 1928

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-611-307-16

Query Match 33.3%; Score 32; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EHCNENDVTT 11  
Db 2 EHCNENDVTT 11

RESULT 4

US-08-758-621-21  
Sequence 21, Application US/08758621  
Patent No. 5846821

GENERAL INFORMATION:

APPLICANT: Gueriot, Mary Lou, and Elide, David J.

TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE &amp; COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/758,621

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/018,578

FILING DATE: 29-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Silveri, Jean M.

REGISTRATION NUMBER: 39,030

REFERENCE/DOCKET NUMBER: DCI-099CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-758-621-21

Query Match 31.2%; Score 30; DB 2; Length 16;  
Best Local Similarity 45.5%; Pred. No. 57;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTTRLRENE 17  
| | | | : : : :  
Db 3 NDVTLPKEDD 13

RESULT 5  
US-09-107-858-21  
; Sequence 21, Application US/09107858  
; Patent No. 6162900  
; GENERAL INFORMATION:  
; APPLICANT: Gueriot, Mary Lou et al.  
; TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR  
; FILE REFERENCE: DCI-099CPDV  
; CURRENT APPLICATION NUMBER: US/09/107,858  
; EARLIER FILING DATE: 1998-06-30  
; EARLIER APPLICATION NUMBER: 08/758,621  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-107-858-21

Query Match 31.2%; Score 30; DB 4; Length 16;  
Best Local Similarity 45.5%; Pred. No. 57;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 NDVTTRLRENE 17  
| | | | : : : :  
Db 3 NDVTLPKEDD 13

RESULT 6  
5496552-8  
; Patent No. 5496552  
; APPLICANT: KUBERASAMPATH, THANGAVEL; RUEGER, DAVID C.  
; TITLE OF INVENTION: OSTEOGENIC DEVICES  
; NUMBER OF SEQUENCES: 25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/268,252  
; FILING DATE: 29-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 103,604  
; FILING DATE: 06-AUG-1993  
; APPLICATION NUMBER: 827,052  
; FILING DATE: 28-JAN-1992  
; APPLICATION NUMBER: 579,865  
; FILING DATE: 07-SEP-1990  
; APPLICATION NUMBER: 179,406  
; FILING DATE: 08-APR-1988  
; SEQ ID NO: 8  
; LENGTH: 11  
5496552-8

Query Match 30.2%; Score 29; DB 6; Length 11;

Best Local Similarity 50.0%; Pred. No. 55;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FEHCNFD 8  
| | | : :  
Db 1 FLHCQFSE 8

RESULT 7  
US-08-460-309-23  
; Sequence 23, Application US/08460309  
; Patent No. 5837496  
; GENERAL INFORMATION:  
; APPLICANT: Leivo, Ilmo  
; APPLICANT: Leivo, Ilmo  
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,309  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/125,077  
; FILING DATE: 22-SEP-1993  
; APPLICATION NUMBER: US PCT/US 94/10730  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/472,319  
; FILING DATE: 30-JAN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/919,951  
; FILING DATE: 27-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9721  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-460-309-23

Query Match 30.2%; Score 29; DB 2; Length 14;  
Best Local Similarity 38.5%; Pred. No. 72;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 CNFNDVTTRLREN 16  
| : | : | : | :  
Db 1 CSIVDITDQEN 13

RESULT 8  
US-08-125-077-23  
; Sequence 23, Application US/08125077  
; Patent No. 5872231

Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION/DOCKET NUMBER: 31,815  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-125-077-23

Query Match  
Best Local Similarity 30.2%; Score 29; DB 2; Length 14;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 CNEVDVTRLREN 16  
I: | : | | |  
Db 1 CSTVIDTJNOEEN 13

RESULT 9  
5444158-4  
Patent No. 5444158  
APPLICANT: ENGVAL, EVA; SANES, JOSHUA  
TITLE OF INVENTION: MEROSIN, NUCLEIC ACIDS ENCODING,  
FRAGMENTS AND USES THEREOF  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/87,642  
FILING DATE: 08-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 587,689  
FILING DATE: 24-SEP-1990  
APPLICATION NUMBER: 472,319  
FILING DATE: 30-JAN-1990  
SEQ ID NO: 4:  
LENGTH: 14

5444158-4  
Query Match  
Best Local Similarity 30.2%; Score 29; DB 6; Length 14;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 CNEVDVTRLREN 16  
I: | : | | |  
Db 1 CSTVIDTJNOEEN 13

RESULT 10  
US-08-522-326-6  
Sequence 6, Application US/08522326  
Patent No. 5753615  
GENERAL INFORMATION:  
APPLICANT: THORPE, Alan  
APPLICANT: DUVE, Hanne  
TITLE OF INVENTION: NEUROPEPTIDES AND THEIR USE AS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: US  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/522,326  
FILING DATE: 12-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9305120.9  
FILING DATE: 12-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: LUDWIG, STEVEN R.  
REGISTRATION/DOCKET NUMBER: 36,203  
TELEPHONE: (202) 371-2540  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Manduca sexta  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /product= "OTHER"  
PUBLICATION INFORMATION: /note= "pyro-glutamic acid"  
DOCUMENT NUMBER: WO 94/20530 A2  
FILING DATE: 14-MAR-1994  
PUBLICATION DATE: 15-SEP-1994  
US-08-522-326-6

Query Match  
Best Local Similarity 29.2%; Score 28; DB 1; Length 15;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;



OY 1 EHCNFDVT 10  
| | | |  
Db 4 FROCFNPIS 13

## RESULT 11

US-08-556-597-133  
; Sequence 133, Application US/08556597  
; Patent No. 5877155  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Lyle, Vicki A.  
; TITLE OF INVENTION: MIMOTOPES AND ANTI-MIMOTOPES OF  
; TITLE OF INVENTION: HUMAN PLATELET GLYCOPROTEIN Ib/IX  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/556,597  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/406,330  
; FILING DATE: 17-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ. ID NO: 133:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-556-597-133

Query Match 28.6%; Score 27.5; DB 2; Length 10;  
Best Local Similarity 46.2%; Pred. No. 87;  
Matches 6; Conservative 3; Mismatches 1; Indels 3; Gaps 1;  
OY 2 EHCNFDVT 14  
| | | |  
Db 1 KHCS---IHTRLR 10

## RESULT 12

US-08-685-357B-12  
; Sequence 12, Application US/08685357B  
; Patent No. 5854216  
; GENERAL INFORMATION:  
; APPLICANT: GAUDREAU, Pierrette  
; TITLE OF INVENTION: MARKER FOR GROWTH HORMONE-RELEASING  
; TITLE OF INVENTION: FACTOR RECEPTORS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kevin M. Farrell, P.C.  
; STREET: 12 Riverwood Drive - P.O. Box 999  
; CITY: York Harbor

STATE: ME  
COUNTRY: U.S.A.  
ZIP: 03911  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,357B  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/312,244  
; FILING DATE: 23-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FARRELL, Kevin M.  
; REGISTRATION NUMBER: 35,505  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (207) 363-0558  
; TELEFAX: (207) 363-0528  
; INFORMATION FOR SEQ. ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; US-08-685-357B-12

Query Match 28.1%; Score 27; DB 2; Length 12;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 8 DVTTRLRENE 17  
| | | | |  
Db 1 DFTQLRDE 10

RESULT 13  
US-08-952-568-20  
; Sequence 20, Application US/08952568  
; Patent No. 6037321  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Vasoactive Intestinal Peptide  
; NUMBER OF SEQUENCES: 43  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/952,568  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/CA96/00280  
; FILING DATE:  
; INFORMATION FOR SEQ. ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; US-08-952-568-20

Query Match 28.1%; Score 27; DB 3; Length 12;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRRLRE 15  
| | | | |  
Db 1 FTDVTRRLRK 10

RESULT 14  
5194585-1  
; Patent No. 5194585  
; APPLICANT: PAUL, SUDHIR; POWELL, MICHAEL J.  
; MASSEY, RICHARD J.  
; TITLE OF INVENTION: INHIBITORS OF CATALYTIC ANTIBODIES  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/486,594  
; FILING DATE: 28-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 343,081  
; FILING DATE: 25-APR-1989  
; SEQ ID NO: 1  
; LENGTH: 16  
5194585-1

Query Match 28.1%; Score 27; DB 6; Length 16;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRRLRE 15  
| | | | |  
Db 6 FTDVTRRLRK 15

RESULT 15  
5194585-3  
; Patent No. 5194585  
; APPLICANT: PAUL, SUDHIR; POWELL, MICHAEL J.  
; MASSEY, RICHARD J.  
; TITLE OF INVENTION: INHIBITORS OF CATALYTIC ANTIBODIES  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/486,594  
; FILING DATE: 28-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 343,081  
; FILING DATE: 25-APR-1989  
; SEQ ID NO: 3  
; LENGTH: 16  
5194585-3

Query Match 28.1%; Score 27; DB 6; Length 16;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 FNDVTRRLRE 15  
| | | | |  
Db 6 FTDVTRRLRK 15

RESULT 16  
US-08-218-025A-105  
; Sequence 105; Application US/08218025A  
; Patent No. 5556744  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, David B.  
; APPLICANT: Ugen, Kenneth E.  
; APPLICANT: Williams, William V.  
; TITLE OF INVENTION: Methods and Compositions for Diagnosing  
; NUMBER OF SEQUENCES: and Treating Certain HIV Infected Patients  
; CORRESPONDENCE ADDRESS: 197

ADDRESSEE: Howson and Howson  
STREET: P.O. Box 457, 321 No. 55567441sttown Road  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,025A  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/891,451  
FILING DATE: 29-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST33A  
TELEPHONE: (215) 540-9206  
TELEFAX: (215) 540-5818  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-218-025A-105

Query Match 27.6%; Score 26.5; DB 1; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 4 CNDVTRRLR 14  
| | | | |  
Db 1 CSFM-ISTSR 10

RESULT 17  
US-09-224-785-22  
; Sequence 22; Application US/09224785A  
; Patent No. 6197526  
; GENERAL INFORMATION:  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Yu, Jinan  
; APPLICANT: Kelley, Brian D.  
; APPLICANT: Deetz, Jeffrey S.  
; APPLICANT: Booth, James E.  
; TITLE OF INVENTION: Binding Molecules for Human Factor VIII and Factor  
; FILE REFERENCE: Dyax-008 US sequence listing  
; CURRENT APPLICATION NUMBER: US/09/224,785A  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 11  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-224-785-22

Query Match 27.1%; Score 26; DB 4; Length 11;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 PEHCH 5  
1:11:  
Db 7 FDHCH 11

## RESULT 18

US-09-224-785-25  
Sequence 25, Application US/09224785A  
Patent No. 6197526  
GENERAL INFORMATION:  
APPLICANT: Potter, M. Daniel  
APPLICANT: Yu, Jinan  
APPLICANT: Kelley, Brian D  
APPLICANT: Deetz, Jeffrey S  
TITLE OF INVENTION: Binding Molecules for Human Factor VIII and Factor  
FILE REFERENCE: Dvax-008 US sequence listing  
CURRENT APPLICATION NUMBER: US/09/224,785A  
CURRENT FILING DATE: 1999-01-04  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 25  
LENGTH: 11  
TYPE: PRP  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-224-785-25

Query Match 27.1%; Score 26; DB 4; Length 11;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 PEHCH 5  
1:11:  
Db 7 FDHCH 11

RESULT 19  
US-07-778-233B-33  
Sequence 33, Application US/07778233B  
Patent No. 5270170  
GENERAL INFORMATION:  
APPLICANT: Schatz, Peter J.  
APPLICANT: Cull, Willard G.  
TITLE OF INVENTION: Peptide Library and Screening Method  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/778,233B  
FILING DATE: 19911016  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-50  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-326-2400  
FAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 7 4 1.1  
US-07-778-233B-33

Query Match 27.1%; Score 26; DB 1; Length 12;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 VTPRLRENE 17  
1:1:1:1:  
Db 1 VTPRLRSNK 9

## RESULT 20

US-07-963-321-33  
Sequence 33, Application US/07963321  
Patent No. 538665  
GENERAL INFORMATION:  
APPLICANT: Schatz, Peter J.  
APPLICANT: Cull, Willard G.  
APPLICANT: Miller, Jeff F.  
APPLICANT: Stemmer, Willem P.C.  
TITLE OF INVENTION: Peptide Library and Screening Method  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/963,321  
FILING DATE: 19921015  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/778,223  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11509-50-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
FAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 7 4 1.1  
US-07-963-321-33

Query Match 27.1%; Score 26; DB 1; Length 12;

Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 VTTRPRENE 17  
| | | | |  
Db 1 VTORVRNKK 9

## RESULT 21

US-08-290-641-33  
; Sequence 33, Application US/08290641  
; Patent No. 5498530  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, William P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/290,641  
; FILING DATE: 15-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,321  
; FILING DATE: 15-OCT-1992  
; APPLICATION NUMBER: US 07/778,223  
; FILING DATE: 16-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11509-50-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 7 4 1.1  
US-08-290-641-33

Query Match 27.1%; Score 26; DB 1; Length 12;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 VTTRPRENE 17  
| | | | |  
Db 1 VTORVRNKK 9

## RESULT 22

US-08-548-540-33  
; Sequence 33, Application US/08548540  
; Patent No. 5733731  
; GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.  
APPLICANT: Cull, Millard G.  
APPLICANT: Miller, Jeff F.  
APPLICANT: Stemmer, William P.C.  
TITLE OF INVENTION: Peptide Library and Screening Method  
NUMBER OF SEQUENCES: 162  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/548,540  
FILING DATE: 26-OCT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/290,641  
FILING DATE: 15-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,321  
FILING DATE: 15-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 16528J-0012400S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 7 4 1.1  
US-08-548-540-33

Query Match 27.1%; Score 26; DB 1; Length 12;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 VTTRPRENE 17  
| | | | |  
Db 1 VTORVRNKK 9

## RESULT 23

PCT-US96-09809-33  
; Sequence 33, Application PC/TUS9609809  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, William P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 162  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California

COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09809  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/548,540  
FILING DATE: 26-OCT-1995  
APPLICATION NUMBER: US 08/290,641  
FILING DATE: 15-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,321  
FILING DATE: 15-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 16528J-001240US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 741.1  
PCT-US96-09809-33

Query Match 27.1%; Score 26; DB 5; Length 12;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 VTTTLLRENE 17  
| | | | |  
DB 1 VTFQVRVSNK 9

RESULT 24  
US-08-413-708B-5  
Sequence 5, Application US/08413708B  
Patent No. 5972883  
GENERAL INFORMATION:  
APPLICANT: GOZES, Iliana  
APPLICANT: FRIDKIN, Matlyahu  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING  
TITLE OF INVENTION: VIP, OR ANALOGUES DERIVATIVES AND FRAGMENTS THEREOF  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEWMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,708B  
FILING DATE: 30-MAR-1995  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/207,671  
FILING DATE: 09-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 105061  
FILING DATE: 16-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: GOZES-3A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-413-708B-5

Query Match 27.1%; Score 26; DB 2; Length 14;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 6 FNDVTRRLR 14  
| | | | |  
DB 6 FTDNYTRRLR 14

RESULT 25  
US-09-192-048-24  
Sequence 24, Application US/09192048  
Patent No. 6063592  
GENERAL INFORMATION:  
APPLICANT: Lee, Soohae  
TITLE OF INVENTION: Kell Protein Proteolytic Activity  
FILE REFERENCE: Sequence 1-27  
Patent No. 6063592  
CURRENT APPLICATION NUMBER: US/09/192,048  
CURRENT FILING DATE: 1998-11-13  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 24  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-192-048-24

Query Match 27.1%; Score 26; DB 3; Length 14;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 6 FNDVTRRLR 14  
| | | | |  
DB 6 FTDNYTRRLR 14

Search completed: April 2, 2002, 09:24:10  
Job time: 79 sec

Tue Apr 2 09:51:48 2002

us-09-020-393b-3\_copy\_42\_58.closed.ra1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:09:12 ; Search time 19.6 Seconds  
(without alignments)  
104.934 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145

Sequence: 1 KCLGYHLDVSLAFSEISVGAERFKDDC 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 5455

Minimum DB seq length: 0

Maximum DB seq length: 27

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 31    | 21.4        | 22     | 2     | T25653      |
| 2          | 28    | 19.3        | 17     | 2     | A60889      |
| 3          | 28    | 19.3        | 22     | 2     | A36389      |
| 4          | 28    | 19.3        | 27     | 2     | A41006      |
| 5          | 27    | 18.6        | 22     | 2     | JP0066      |
| 6          | 27    | 18.6        | 25     | 2     | S07574      |
| 7          | 26.5  | 18.3        | 20     | 2     | S72501      |
| 8          | 26    | 17.9        | 15     | 2     | S08301      |
| 9          | 26    | 17.9        | 18     | 2     | A61392      |
| 10         | 26    | 17.9        | 21     | 2     | I65270      |
| 11         | 26    | 17.9        | 24     | 1     | B32252      |
| 12         | 26    | 17.9        | 24     | 1     | T46622      |
| 13         | 26    | 17.9        | 25     | 2     | A48810      |
| 14         | 25    | 17.2        | 20     | 2     | A60295      |
| 15         | 25    | 17.2        | 27     | 2     | T37261      |
| 16         | 24    | 16.6        | 17     | 2     | S71864      |
| 17         | 24    | 16.6        | 17     | 2     | A36727      |
| 18         | 24    | 16.6        | 21     | 2     | S38739      |
| 19         | 24    | 16.6        | 22     | 2     | C46285      |
| 20         | 24    | 16.6        | 22     | 2     | PC2134      |
| 21         | 24    | 16.6        | 23     | 2     | S38738      |
| 22         | 23    | 15.9        | 14     | 2     | PS0371      |
| 23         | 23    | 15.9        | 14     | 2     | B61597      |
| 24         | 23    | 15.9        | 15     | 2     | PA0062      |
| 25         | 23    | 15.9        | 15     | 2     | D56385      |
| 26         | 23    | 15.9        | 15     | 2     | A47628      |
| 27         | 23    | 15.9        | 16     | 2     | B45895      |
| 28         | 23    | 15.9        | 17     | 2     | PH1357      |
| 29         | 23    | 15.9        | 18     | 2     | S36121      |

|    |      |      |    |   |        |                    |
|----|------|------|----|---|--------|--------------------|
| 30 | 23   | 15.9 | 19 | 2 | A38386 | Ca2+-transporting  |
| 31 | 23   | 15.9 | 22 | 2 | A35418 | brain natriuretic  |
| 32 | 23   | 15.9 | 22 | 2 | JT0581 | natriuretic peptid |
| 33 | 23   | 15.9 | 24 | 1 | S58242 | pyrroloquinoline q |
| 34 | 23   | 15.9 | 24 | 2 | S77982 | cytochrome-c oxida |
| 35 | 23   | 15.9 | 25 | 2 | S39360 | CDK inhibitor - mo |
| 36 | 23   | 15.9 | 26 | 2 | S59906 | gamma-glutamyl tra |
| 37 | 23   | 15.9 | 26 | 2 | D53507 | annexin V - rat (f |
| 38 | 23   | 15.9 | 27 | 2 | JC1081 | brain natriuretic  |
| 39 | 22.5 | 15.5 | 26 | 4 | T07533 | probable photosyst |
| 40 | 22   | 15.2 | 11 | 2 | A35594 | buccalin - Callfor |
| 41 | 22   | 15.2 | 15 | 2 | S66215 | cartilage oligomer |
| 42 | 22   | 15.2 | 20 | 2 | S23981 | outer layer protei |
| 43 | 22   | 15.2 | 20 | 2 | A61526 | major milk gland p |
| 44 | 22   | 15.2 | 21 | 2 | F64121 | hypothetical prote |
| 45 | 22   | 15.2 | 23 | 1 | S20453 | pyrroloquinoline q |

## ALIGNMENTS

RESULT 1  
T25653  
hypothetical protein C47C12.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T25653  
R:Connell, M.  
submitted to the EMBL Data Library, August 1996  
A:Description: The sequence of C. elegans cosmid C47C12.  
A:Reference number: Z20062  
A:Accession: T25653  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-22 <CON>  
A:Cross-references: EMBL:U67951; PIDB:AB07573.1; GSPDB:GN00028; CESP:C47C12.5  
A:Experimental source: strain Bristol N2; clone C47C12  
C:Genetics:  
A:Gene: CESP:C47C12.5  
A:Map position: X  
A:Introns: 21/1

Query Match 21.4%; Score 31; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
Matches 9; Conservative 2; Mismatches 5; Indels 2; Gaps 1;  
OY 6 HLDV--SLAFSEISVGA 21  
| : | | | | : |  
Db 4 HSDIGSLQFSRISLKNE 21  
RESULT 2  
A60889  
olfactory glycoprotein RB-8 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: A60889  
R:Schwab, J.E.; Gottlieb, D.I.  
J. Neurosci. 8, 3470-3480, 1988  
A:Title: Purification and characterization of an antigen that is spatially segregated  
A:Reference number: A60889; MUID:89010968  
A:Accession: A60889  
A:Molecule type: protein  
A:Residues: 1-17 <SCH>  
A:Note: this protein was purified from whole brain  
C:Comment: The monoclonal antibody RB-8 binds this integral membrane glycoprotein on  
C:Keywords: glycoprotein; membrane protein; olfaction

Query Match 19.3%; Score 28; DB 2; Length 17;  
Best Local Similarity 54.5%; Pred. No. 5.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VSLAFSEISVG 19  
 :||: |||||  
 Db 6 ISLSKVELSVG 16

RESULT 3  
 A36399  
 C-type natriuretic peptide - frog  
 C:Species: Ranidae gen. sp. (frog)  
 C>Date: 01-Feb-1991 #sequence\_revision 01-Feb-1991 #text\_change 09-Dec-1994  
 C:Accession: A36399  
 R:Yoshihara, A.; Kozawa, H.; Minamino, N.; Kangawa, K.; Matsuo, H.  
 Biochem. Biophys. Res. Commun. 173, 591-598, 1990  
 A:Title: Isolation and sequence determination of frog C-type natriuretic peptide.  
 A:Reference number: A36399; MUID:91083642  
 A:Accession: A36399  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-22 <YOS>  
 C:Superfamily: natriuretic peptide A precursor

Query Match 19.3%; Score 28; DB 2; Length 22;  
 Best Local Similarity 53.8%; Pred. No. 7.3e+02;  
 Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 CLGYHDLVSLAFS 14  
 | | | | |  
 Db 6 CFCVGLDRIKAFS 18

RESULT 4  
 A41006  
 LDL receptor - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C>Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 23-Mar-1993  
 C:Accession: A41006  
 R:Barber, D.L.; Sanders, E.J.; Aebersold, R.; Schneider, W.J.  
 J. Biol. Chem. 266, 18761-18770, 1991  
 A:Title: The receptor for yolk lipoprotein deposition in the chicken oocyte.  
 A:Reference number: A41006; MUID:92011658  
 A:Accession: A41006  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-27 <BAK>

Query Match 19.3%; Score 28; DB 2; Length 27;  
 Best Local Similarity 31.8%; Pred. No. 9.1e+02;  
 Matches 7; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
 QY 5 YHLDVSLAFSEISVGAEFNKDD 26  
 | | | | | : | | |  
 Db 3 YWTDSSAIFSAISIDTSMNFDN 24

RESULT 5  
 JP0066  
 ribosomal protein L30 - Nocardia asteroides (fragment)  
 C:Species: Nocardia asteroides  
 C>Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 28-Oct-1994  
 C:Accession: JP0066  
 R:Ochi, K.  
 submitted to JPIPD, February 1994  
 A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal p  
 A:Reference number: JP0066  
 A:Accession: JP0066  
 A:Molecule type: protein  
 A:Residues: 1-22 <OCH>  
 C:Keywords: protein biosynthesis; ribosome

Query Match 18.6%; Score 27; DB 2; Length 22;  
 Best Local Similarity 26.3%; Pred. No. 1e+03;  
 Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 QY 8 DVSLAFSEISVGAEFNKDD 26  
 | : : : | : | : |  
 Db 2 DLKVTQIKSTIGAKANKD 20

RESULT 6  
 S07574  
 malate dehydrogenase (EC 1.1.1.37) - Phenyllobacterium immobile (fragment)  
 C:Species: Phenyllobacterium immobile  
 C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 02-Aug-1994  
 C:Accession: S07574  
 R:Rommel, T.O.; Hurd, H.K.; Speth, A.R.; Lingens, F.  
 Biol. Chem. Hoppe-Seyler 370, 763-768, 1989  
 A:Title: Purification and N-terminal amino-acid sequences of bacterial malate dehydr  
 A:Reference number: S04956; MUID:89374824  
 A:Accession: S07574  
 A:Molecule type: protein  
 A:Residues: 1-25 <ROM>  
 C:Superfamily: L-lactate dehydrogenase  
 C:Keywords: oxidoreductase; tricarboxylic acid cycle

Query Match 18.6%; Score 27; DB 2; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 LGYHL 7  
 :|||:  
 Db 16 IGYHL 20

RESULT 7  
 S72501  
 protein kinase C inhibitor - human (fragment)  
 N:Alternate names: histidine triad nucleotide-binding protein  
 C:Species: Homo sapiens (man)  
 C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
 C:Accession: S72501; S62623  
 R:Maines, M.D.; Trakshel, G.M.  
 Arch. Biochem. Biophys. 300, 320-326, 1993  
 A:Title: Purification and characterization of human biliverdin reductase.  
 A:Reference number: S29736; MUID:93143333  
 A:Accession: S72501  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAI>  
 A:Note: this protein was identified as biliverdin reductase; the identification is qu  
 R:Maines, M.D.; Polevoda, B.V.; Huang, T.J.; McCoubrey Jr., W.K.  
 Eur. J. Biochem. 235, 372-381, 1996  
 A:Title: Human biliverdin IX-alpha reductase is a zinc-metalloprotein. Characterizati  
 A:Reference number: S62622; MUID:96202961  
 A:Accession: S62623  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAW>  
 C:Superfamily: protein kinase C inhibitor; histidine triad homology  
 C:Keywords: homodimer; protein kinase inhibitor; zinc

Query Match 18.3%; Score 26.5; DB 2; Length 20;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
 QY 2 CLGYHDLVS 10  
 | | : | | : |  
 Db 1 CLAFH-DIS 8

RESULT 8  
 S08301  
 epidermal growth factor, high molecular weight - rat (fragment)



C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-Sep-1993  
 C:Accession: S08301  
 R:Nexo, E.; Jorgensen, P.E.; Thim, L.; Roepstorff, P.  
 Biochim. Biophys. Acta 1037, 388-393, 1990  
 A:Title: Purification and characterization of a low and a high molecular weight form of  
 A:Reference number: S08288; MUID:90181442  
 A:Accession: S08301  
 A:Molecule type: protein  
 A:Residues: 1-15 <NEX>  
 C:Keywords: growth factor

Query Match 17.9%; Score 26; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. NO. 1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 KDCC 27  
 |||  
 Db 3 KDCC 6

RESULT 9  
 A61392  
 brain-associated small cell lung cancer antigen - human (fragment)  
 N:Alternate names: BASCA  
 C:Species: Homo sapiens (man)  
 C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Sep-1994  
 C:Accession: A61392 - sequence\_revision 09-Sep-1994 #text\_change 09-Sep-1994  
 R:Umezawa, Y.; Kuge, S.; Kikyo, N.; Shirai, T.; Watanabe, J.; Fujiwara, M.; Okabe, T.  
 Jpn. J. Clin. Oncol. 21, 251-255, 1991  
 A:Title: Identity of brain-associated small cell lung cancer antigen and the CD56 (NKH-1  
 A:Reference number: A61392; MUID:92046737  
 A:Accession: A61392  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-18 <DME>

Query Match 17.9%; Score 26; DB 2; Length 18;  
 Best Local Similarity 46.2%; Pred. NO. 1.2e+03;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 7 LDVSLAFSEISVG 19  
 |||  
 Db 3 VDIVPSGSEISVG 15

RESULT 10  
 I65270  
 collagen alpha 1(I) chain - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 13-Aug-1999  
 C:Accession: I65270  
 R:Genovese, C.; Rowe, D.; Kream, B.  
 Biochemistry 23, 6210-6216, 1984  
 A:Title: Construction of DNA sequences complementary to rat alpha-1 and alpha-2 collagen  
 A:Reference number: I52392; MUID:85122694  
 A:Accession: I65270  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-21 <RES>  
 A:Cross-references: GB:M12200; NID:g203191; PID:AAA0835.1; PID:g203195  
 C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 17.9%; Score 26; DB 2; Length 21;  
 Best Local Similarity 33.3%; Pred. NO. 1.4e+03;  
 Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 5 YHLDVSLAFSEISVG 19  
 |||  
 Db 7 YHCKNSIAYLDEMG 21

RESULT 11  
 B32252  
 pyrroloquinoline quinone precursor - Acinetobacter calcoaceticus

N:Alternate names: pqg gene IV protein  
 C:Species: Acinetobacter calcoaceticus  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
 C:Accession: B32252  
 R:Goosen, N.; Horsman, H.P.A.; Huinen, R.G.M.; van de Putte, P.  
 J. Bacteriol. 171, 447-455, 1989

A:Title: Acinetobacter calcoaceticus genes involved in biosynthesis of the coenzyme P  
 A:Reference number: A32252; MUID:89123056

A:Accession: B32252

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-24 <GOO>

A:Cross-references: GB:X06452; NID:g38740; PID:g130796

A:Note: this ORF is not annotated in Genbank entry ACPQ05YN, release 116.0

R:Goosen, N.; Huinen, R.G.; van de Putte, P.

J. Bacteriol. 174, 1426-1427, 1992

A:Title: A 24-amino-acid polypeptide is essential for the biosynthesis of the coenzyme

A:Reference number: A59183; MUID:92138642

A:Contents: annotation

C:Genetics:

A:Gene: pqgA

C:Superfamily: pyrroloquinoline quinone precursor pqgA

C:Keywords: quinoprotein

F:16/20/Product: pyrroloquinoline quinone #status predicted <MAN>

F:16/20/Cross-Link: pyrroloquinoline quinone (Glu, Tyr) #status predicted

Query Match 17.9%; Score 26; DB 1; Length 24;  
 Best Local Similarity 40.0%; Pred. NO. 1.6e+03;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 12 AFSEISVGAE 21  
 |||  
 Db 7 AFTDLRIGFE 16

RESULT 12  
 T46622  
 hypothetical protein cl - lobliolly pine  
 C:Species: Pinus taeda (loblolly pine)  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
 C:Accession: T46622  
 R:Chang, S.; Puryear, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.  
 submitted to the EMBL Data Library, July 1995  
 A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is  
 A:Reference number: Z23105  
 A:Accession: T46622  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-24 <CHA>  
 A:Cross-references: EMBL:U01309; NID:g974285; PID:g974287  
 A:Experimental source: strain s6PT2x56PT3; 8 month seedlings

Query Match 17.9%; Score 26; DB 2; Length 24;  
 Best Local Similarity 41.7%; Pred. NO. 1.6e+03;  
 Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 CLGYHLDVSLAF 13  
 |||  
 Db 11 CGGFOPDCLICF 22

RESULT 13  
 A48810  
 fibrinogen B beta subunit - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C:Accession: A48810  
 R:Roberts, L.R.; Nichols, L.A.; Holland, L.J.  
 Biochemistry 32, 11627-11637, 1993  
 A:Title: Transcriptional regulation of the Xenopus laevis beta fibrinogen subunit gene  
 A:Reference number: A48810; MUID:94032285  
 A:Accession: A48810  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-25 <ROB>  
 A:Cross-references: GB:U05035; GB:S66373; NID:9450950; PIDN:AAA60463.1; PID:9450951  
 A:Note: sequence extracted from NCBI backbone (NCBIN:138880, NCBIPI:138881)

Query Match 17.9%; Score 26; DB 2; Length 25;  
 Best Local Similarity 23.5%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 10 SLAFSISVGAEPNKDD 26  
 :| | : : : : :|  
 Db 8 ALCVSAWCSDDYEDD 24

RESULT 14  
 A60295  
 Apolipophorin III - house cricket (fragment)  
 C:Species: Acheta domestica (house cricket)  
 C:Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 18-Jun-1993  
 A:Accession: A60295  
 R:Strobel, L.M.; Kanost, M.R.; Ziegler, R.; Wells, M.A.  
 Insect Biochem. 20, 859-863, 1990  
 A:Title: Adipokinetic hormone causes formation of a low density lipophorin in the house  
 A:Reference number: A60295  
 A:Accession: A60295  
 A:Molecule type: protein  
 A:Residues: 1-20 <STR>  
 C:Comment: This protein, a small, water-soluble apolipoprotein, is thought to increase t  
 ts a residual pathway for flight-related lipid transport.  
 C:Keywords: hemolymph; lipid transport

Query Match 17.2%; Score 25; DB 2; Length 20;  
 Best Local Similarity 57.1%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 SVGAEPN 23  
 :| | : : : : :|  
 Db 4 TTGADFN 10

RESULT 15  
 I37261  
 complement receptor - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
 A:Accession: I37261  
 R:Bitenbach, M.; Tong, X.; Bradbury, L.E.; Tedder, T.F.; Kieff, E.  
 J. Exp. Med. 176, 1405-1414, 1992  
 A:Title: Characterization of an Epstein-Barr virus receptor on human epithelial cells.  
 A:Reference number: I37261; MUID:93018869  
 A:Accession: I37261  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-27 <RES>  
 A:Cross-references: EMBL:X68990; NID:93928195; PIDN:CAA48779.1; PID:9395334  
 C:Genetics:  
 A:Gene: GDB:CR2  
 A:Cross-references: GDB:119802; OMIM:120650  
 A:Map position: Iq32-Iq32

Query Match 17.2%; Score 25; DB 2; Length 27;  
 Best Local Similarity 33.3%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 CLGYHDLVSIAF 13  
 | | : | :|  
 Db 6 CAGHYLRRLSM 17

RESULT 16  
 S71864  
 glutathione transferase (EC 2.5.1.18) class alpha 6a - pig (fragment)  
 N:Alternate names: glutathione S-transferase class alpha 6a  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 07-May-1999  
 A:Accession: S71864  
 R:Rouini, P.; Anglade, P.; Debrauwer, L.; Tulliez, J.  
 Biochem. J. 317, 879-884, 1996  
 A:Title: Characterization of pig liver glutathione S-transferases using HPLC-electros  
 A:Reference number: S71864; MUID:96332484  
 A:Accession: S71864  
 A:Molecule type: protein  
 A:Residues: 1-17 <ROU>  
 A:Experimental source: liver; cytosolic  
 C:Comment: At least five species-independent classes of cytosolic glutathione transfer  
 s mitochondrial form are known.  
 C:Function: dimer  
 A:Description: catalyzes the nucleophilic conjugation of intracellular glutathione to  
 A:Pathway: detoxification; xenobiotics metabolism  
 A:Note: Increased hydrophilicity of GSH-conjugates facilitates their further metabol  
 es of damage  
 C:Superfamily: glutathione transferase  
 C:Keywords: dimer; transferase

Query Match 16.6%; Score 24; DB 2; Length 17;  
 Best Local Similarity 35.7%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 8 DVSLAFSEISVGA 21  
 | | | : :|  
 Db 1 DGSLLFQVDMYTE 14

RESULT 17  
 A36727  
 cytochrome c551 - Methylinomas sp. (fragment)  
 C:Species: Methylinomas sp.  
 C:Date: 19-Apr-1991 #sequence\_revision 19-Apr-1991 #text\_change 18-Jun-1993  
 A:Accession: A36727  
 R:DiSpirito, A.A.; Lipscomb, J.D.; Lidstrom, M.E.  
 J. Bacteriol. 172, 5360-5367, 1990  
 A:Title: Soluble cytochromes from the marine methanotroph Methylinomas sp. strain A4.  
 A:Reference number: A36727; MUID:90366596  
 A:Accession: A36727  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-17 <DIS>

Query Match 16.6%; Score 24; DB 2; Length 17;  
 Best Local Similarity 60.0%; Pred. No. 2.3e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CLGYH 6  
 | | :|  
 Db 13 CSGFH 17

RESULT 18  
 S38739  
 lipid transfer protein S2 - spinach (fragment)  
 C:Species: Spinacia oleracea (spinach)  
 C:Date: 08-Jun-1994 #sequence\_revision 19-Jan-1996 #text\_change 03-May-1996  
 A:Accession: S38739

R;Segura, A.; Moreno, M.; Garcia-Olmedo, F.  
FEBS Lett. 332, 243-246, 1993  
A:Title: Purification and antipathogenic activity of lipid transfer proteins (LTPs) from  
A:Reference number: S38736; MUID:94009709  
A:Accession: S38739  
A:Molecule type: protein  
A:Residues: 1-21 <SEG>  
C:Superfamily: phospholipid transfer protein

Query Match 16.6%; Score 24; DB 2; Length 21;  
Best Local Similarity 75.0%; Pred. No. 2.9e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CLGY 5  
|:|:  
Db 14 CIGY 17

RESULT 19  
C46285  
formaldehyde dehydrogenase (glutathione) (EC 1.2.1.1) class III low activity form - Atla  
C:Species: Gadus morhua (Atlantic cod)  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1996  
C:Accession: C46285  
R:Danielsson, O.; Jornvall, H.  
Proc. Natl. Acad. Sci. U.S.A. 89, 9247-9251, 1992  
A:Title: "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutath  
A:Reference number: A46285; MUID:93028441  
A:Accession: C46285  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <DAN>  
A:Note: sequence extracted from NCBI backbone (NCBP:116269)  
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology  
C:Keywords: NAD; oxidoreductase

Query Match 16.6%; Score 24; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 20 AERNKDC 27  
|:|:  
Db 11 AEEGATEC 18

RESULT 20  
PC2134  
maltose transport protein - yeast (Saccharomyces cerevisiae) (fragment)  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
C:Accession: PC2134  
R:Van den Broek, P.J.A.; Van Leeuwen, C.C.M.; Weusthuis, R.A.; Postma, E.; Van Dijken, J  
Bloem. Biophys. Res. Commun. 200, 45-51, 1994  
A:Title: Identification of the maltose transport protein of Saccharomyces cerevisiae.  
A:Reference number: PC2133; MUID:94220124  
A:Accession: PC2134  
A:Molecule type: protein  
A:Residues: 1-22 <VAN>  
A:Experimental source: plasma membrane  
C:Superfamily: maltose transport protein MAL61  
C:Keywords: membrane protein

Query Match 16.6%; Score 24; DB 2; Length 22;  
Best Local Similarity 35.7%; Pred. No. 3.1e+03;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 13 FSEISVGAENKDD 26  
|:|:  
Db 7 FAAMAAAEINVKD 20

RESULT 21  
S38738  
lipid transfer protein S1 - spinach (fragment)  
C:Species: Spinacia oleracea (spinach)  
C:Date: 08-Jun-1994 #sequence\_revision 19-Jan-1996 #text\_change 03-May-1996  
C:Accession: S38738  
R;Segura, A.; Moreno, M.; Garcia-Olmedo, F.  
FEBS Lett. 332, 243-246, 1993  
A:Title: Purification and antipathogenic activity of lipid transfer proteins (LTPs) f  
A:Reference number: S38736; MUID:94009709  
A:Accession: S38738  
A:Molecule type: protein  
A:Residues: 1-23 <SEG>  
C:Superfamily: phospholipid transfer protein

Query Match 16.6%; Score 24; DB 2; Length 23;  
Best Local Similarity 75.0%; Pred. No. 3.2e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CLGY 5  
|:|:  
Db 14 CIGY 17

RESULT 22  
PS0371  
hypothetical protein (psac region) - Synechococcus sp. (fragment)  
C:Species: Synechococcus sp.  
C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
C:Accession: PS0371  
R:Rhee, E.; Stirewalt, V.L.; Gasparich, G.E.; Bryant, D.A.  
Gene 112, 123-128, 1992  
A:Title: The psac genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa: cloning  
A:Reference number: JS0694; MUID:92201692  
A:Accession: PS0371  
A:Molecule type: DNA  
A:Residues: 1-14 <RRH>  
A:Cross-references: GB:86238; NID:9154574; PIDN:AAA27351.1; PID:9552030

Query Match 15.9%; Score 23; DB 2; Length 14;  
Best Local Similarity 57.1%; Pred. No. 2.7e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYHLDVS 10  
|:|:  
Db 3 GFKLDVT 9

RESULT 23  
B61597  
cytochrome P450 AL-2 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
C:Accession: B61597  
R:Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.  
Drug Metab. Dispos. 19, 291-297, 1991  
A:Title: Purification and aminopyrine monooxygenase activity of liver microsomal cyto  
A:Reference number: A61597; MUID:91292910  
A:Accession: B61597  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <SHI>

Query Match 15.9%; Score 23; DB 2; Length 14;  
Best Local Similarity 38.5%; Pred. No. 2.7e+03;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVG 19  
|:|:  
Db 7 LDVSLAFSEISVG 19

Db 1 MDLLSLSPFLVC 13

RESULT 24

PA0062  
 fumarate hydratase (EC 4.2.1.2) - fungus (Fusarium sporotrichioides) (fragment)  
 C:Species: Fusarium sporotrichioides  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
 C:Accession: PA0062  
 R:Chow, L.P.; Fukaya, N.; Suglura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.  
 submitted to JIPID, October 1994  
 A:Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides  
 A:Reference number: PA0051  
 A:Accession: PA0062  
 A:Molecule type: protein  
 A:Residues: 1-15 <CHO>  
 C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match

Best Local Similarity 15.9%; Score 23; DB 2; Length 15;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 12 AFSEISV 18

Db 9 AFGEIOV 15

RESULT 25

D56385  
 nitrophorin 4 - Rhodnius prolixus (fragment)  
 N:Alternate names: nitric oxide-carry protein NP4  
 C:Species: Rhodnius prolixus  
 C:Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 17-Mar-1999  
 C:Accession: D56385  
 R:Champagne, D.E.; Nussenzweig, R.H.; Ribeiro, J.M.C.  
 J. Biol. Chem. 270, 8691-8695, 1995  
 A:Title: Purification, partial characterization, and cloning of nitric oxide-carrying heme  
 A:Reference number: A56385; MUID:95238361  
 A:Accession: D56385  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <CHA>  
 C:Keywords: heme; salivary gland; vasodilator

Query Match

Best Local Similarity 15.9%; Score 23; DB 2; Length 15;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 FNKND 25

Db 12 FNKND 15

Search completed: January 6, 2002, 09:13:53  
 Job time: 281 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:13:07 ; Search time 12.95 Seconds  
(without alignments)

76.444 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1552

Minimum DB seq length: 0

Maximum DB seq length: 27

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description         |
|------------|-------|-------------|--------|--------------|---------------------|
| 1          | 27    | 18.6        | 25     | 1 MDH_PHEIM  | P19980 phenylloact  |
| 2          | 26    | 17.9        | 24     | 1 PQA_AICIA  | P27532 acinetobact  |
| 3          | 25    | 17.2        | 20     | 1 TPX_CLOPA  | P81361 clostridium  |
| 4          | 24    | 16.6        | 15     | 1 RML2_YEAST | P36522 saccharomyc  |
| 5          | 23    | 15.9        | 12     | 1 XYLA_STRVN | P14405 streptomyc   |
| 6          | 23    | 15.9        | 13     | 1 BP37_LEUMA | P81754 leucophaea   |
| 7          | 23    | 15.9        | 18     | 1 AGI_EUPCH  | P31888 euphorbia c  |
| 8          | 23    | 15.9        | 22     | 1 ANFC_CHICK | P21805 gallus gall  |
| 9          | 23    | 15.9        | 23     | 1 FMK7_PSEAE | O53391 pseudomonas  |
| 10         | 23    | 15.9        | 24     | 1 COXC_THUOB | P80973 thunnus obe  |
| 11         | 23    | 15.9        | 24     | 1 PQA_PSEFL  | P55171 pseudomonas  |
| 12         | 23    | 15.9        | 27     | 1 KT39_PICKL | P80326 picchia kluy |
| 13         | 22    | 15.2        | 20     | 1 VMO2_CHICK | O99849 gallus gall  |
| 14         | 22    | 15.2        | 21     | 1 YD90_HAEIN | P45194 haemophilus  |
| 15         | 22    | 15.2        | 23     | 1 PQA_KLEPN  | P27503 klebsiella   |
| 16         | 22    | 15.2        | 25     | 1 PETM_SKECO | O96810 skeletonema  |
| 17         | 21.5  | 14.8        | 27     | 1 CX7A_CONTE | P24160 conus texti  |
| 18         | 21    | 14.5        | 10     | 1 TK14_LOCMT | P30250 locusta mig  |
| 19         | 21    | 14.5        | 19     | 1 FIBR_SHEEP | P14451 ovis aries   |
| 20         | 21    | 14.5        | 22     | 1 VI01_VACCV | P16714 vaccinia vi  |
| 21         | 21    | 14.5        | 22     | 1 RL10_KLEPN | P41190 klebsiella   |
| 22         | 21    | 14.5        | 23     | 1 RL10_ENTCL | O47608 enterobacte  |
| 23         | 21    | 14.5        | 24     | 1 CAMT_PINPS | P81081 pinus pinas  |
| 24         | 21    | 14.5        | 27     | 1 ANF_ANGJA  | P18144 anguilla ja  |
| 25         | 21    | 14.5        | 27     | 1 CD55_SHEEP | P58020 ovis aries   |
| 26         | 21    | 14.5        | 27     | 1 CXK7_CONPU | P56633 conus purpu  |
| 27         | 20.5  | 14.1        | 10     | 1 YMCD_EMENT | P03884 emeritella   |
| 28         | 20    | 13.8        | 17     | 1 SP34_DICMU | P81545 dicyostella  |
| 29         | 20    | 13.8        | 17     | 1 NU4M_TRIRU | O36834 trichophyto  |
| 30         | 20    | 13.8        | 18     | 1 AGI_EUPMA  | P33889 euphorbia m  |
| 31         | 20    | 13.8        | 21     | 1 UP04_CAEEL | P55955 caenorhabdi  |
| 32         | 20    | 13.8        | 24     | 1 ATPB_MICLU | P80286 micrococcus  |
| 33         | 20    | 13.8        | 24     | 1 PL13_PLETR | P36987 plectreury   |

|    |      |      |    |              |                    |
|----|------|------|----|--------------|--------------------|
| 34 | 20   | 13.8 | 25 | 1 GBX1_MOUSE | P82976 mus musculu |
| 35 | 20   | 13.8 | 27 | 1 DBH1_BIFLO | P17615 bifidobacte |
| 36 | 19.5 | 13.4 | 21 | 1 BTX_ATRBI  | P80163 atricapsis  |
| 37 | 19.5 | 13.4 | 27 | 1 RL22_HALCU | P05973 halobacteri |
| 38 | 19   | 13.1 | 10 | 1 TRP8_LEUMA | P81740 leucophaea  |
| 39 | 19   | 13.1 | 12 | 1 V25K_WSSV  | P82004 white spot  |
| 40 | 19   | 13.1 | 14 | 1 MAST_POLJA | P01517 polistes ja |
| 41 | 19   | 13.1 | 15 | 1 PH2_PERAM  | P82695 periplaneta |
| 42 | 19   | 13.1 | 16 | 1 LE06_BIOGL | P80745 biophalar   |
| 43 | 19   | 13.1 | 16 | 1 RIPK_TRIKI | P16093 trichosan   |
| 44 | 19   | 13.1 | 17 | 1 PH3_PERAM  | P82696 periplaneta |
| 45 | 19   | 13.1 | 17 | 1 PH4_PERAM  | P82697 periplaneta |

## ALIGNMENTS

| RESULT   | ID  | MDH_PHEIM | STANDARD | PRT    | 25 AA. |
|--|---|-----------|----------|--------|--------|
| AC   | P19980  |           |          |        |        |
| DT   | 01-FEB-1991 (Rel. 17, Created)  |           |          |        |        |
| DT   | 01-FEB-1991 (Rel. 17, Last sequence update)                           |           |          |        |        |
| DT   | 30-MAY-2000 (Rel. 39, Last annotation update)                         |           |          |        |        |
| DE   | MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT)                         |           |          |        |        |
| GN   | MDH.  |           |          |        |        |
| OS   | Phenyllobacterium immobile.   |           |          |        |        |
| OC   | Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;       |           |          |        |        |
| OX   | NCBI_TaxID=21;  |           |          |        |        |
| OX   | NCBI_TaxID=21;  |           |          |        |        |
| RN   | [1]   |           |          |        |        |
| RP   | SEQUENCE.   |           |          |        |        |
| RC   | STRAIN=E;   |           |          |        |        |
| RX   | MEDLINE=89374824; PubMed=2775496;                                     |           |          |        |        |
| RA   | Rommel T.O., Hund H.-K., Specht A.R., Lingens F.;                     |           |          |        |        |
| RT   | "Purification and N-terminal amino-acid sequences of bacterial malate |           |          |        |        |
| RT   | dehydrogenases from six actinomycetales strains and from              |           |          |        |        |
| RT   | Phenyllobacterium immobile, strain E."                                |           |          |        |        |
| RL   | Biol. Chem. Hoppe-Seyler 370:763-768(1989).                           |           |          |        |        |
| CC   | -1- CATALYTIC ACTIVITY: L-MALATE + NAD(+) -> OXALOACETATE + NADH.     |           |          |        |        |
| CC   | -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.             |           |          |        |        |
| CC   | PIR: S07574; S07574.  |           |          |        |        |
| DR   | HSSP; P10584; IBDM.   |           |          |        |        |
| DR   | InterPro; IPR001252; MDH_actsite.                                     |           |          |        |        |
| DR   | InterPro; IPR001236; ldh.   |           |          |        |        |
| DR   | Pfam; PF00056; ldb; 1.  |           |          |        |        |
| DR   | PROSITE; PS00068; MDH; PARTIAL.                                       |           |          |        |        |
| KW   | Oxidoreductase; Tricarboxylic acid cycle; NAD.                        |           |          |        |        |
| FT   | NON_TER   |           |          |        |        |
| FT   | NON_TER   |           |          |        |        |
| SO   | SEQUENCE 25 AA; 2626 MW; C8D81E008825845C CRC64;                      |           |          |        |        |
| Query Match  |   |           |          |        |        |
| Best Local Similarity 80.0%; Pred. No. 6.4e+02;            |   |           |          |        |        |
| Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0; |   |           |          |        |        |
| QY   | 3   | LGVHL 7   |          |        |        |
| DB   | 16  | IGVHL 20  |          |        |        |
| RESULT 2   |   |           |          |        |        |
| ID   | PQA_AICIA   | STANDARD; | PRT;     | 24 AA. |        |
| AC   | P27532;   |           |          |        |        |
| DT   | 01-AUG-1992 (Rel. 23, Created)  |           |          |        |        |
| DT   | 01-AUG-1992 (Rel. 23, Last sequence update)                           |           |          |        |        |
| DT   | 01-OCT-1996 (Rel. 34, Last annotation update)                         |           |          |        |        |
| DE   | COENZYME PQQ SYNTHESIS PROTEIN A (COENZYME PQQ SYNTHESIS PROTEIN IV). |           |          |        |        |
| GN   | PQA OR PQQIV.   |           |          |        |        |
| OS   | Acinetobacter calcoaceticus.  |           |          |        |        |
| OC   | Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;           |           |          |        |        |
| OC   | Acinetobacter.  |           |          |        |        |

0X NCBI\_TaxID=471;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LM2 79.41;  
 RX MEDLINE=89123056; PubMed=2536663;  
 RA Goosen N., Horsman H.P.A., Huinen R.G.M., van de Putte P.;  
 RT "Acinetobacter calcoaceticus genes involved in biosynthesis of the  
 coenzyme pyrrolo-quinoline-quinone: nucleotide sequence and  
 expression in *Escherichia coli* K-12.";  
 RL J. Bacteriol. 171:447-455(1989).  
 CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)  
 BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ  
 ARE NECESSARY FOR THE SYNTHESIS.  
 CC -1- SIMILARITY: TO OTHER BACTERIAL PQQ.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X06452; ? NOT ANNOTATED\_CDS.  
 DR PIR: B32252; B32252.  
 KW PQQ.  
 SQ SEQUENCE 24 AA: 2938 MW: 0E671EB9B7EA6838 CRC64;

Query Match 17.9%; Score 26; DB 1; Length 24;  
 Best Local Similarity 40.0%; Pred. No. 8.8e+02;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 12 AFSEISVGA 21  
 DB 7 AFTDIRIGFE 16

RESULT 3  
 ID TPX\_CLOPA STANDARD; PRT; 20 AA.  
 AC P81361;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROBABLE THIOLE PEROXIDASE (EC 1.11.1.-) (CP 42) (FRAGMENT).  
 OS Clostridium pasteurianum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 NCBI\_TaxID=1501;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=W5;  
 RX MEDLINE=98291870; PubMed=9629918;  
 RA Flengsrud R., Skjeldal L.;  
 RT "Two-dimensional gel electrophoresis separation and N-terminal  
 RT sequence analysis of proteins from *Clostridium pasteurianum* W5.";  
 RL Electrophoresis 19:802-806(1998).  
 CC -1- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR  
 H(2)O(2) (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 CC 4.6, ITS MW IS: 20.2 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.  
 DR PROSITE: PS01265; TPX; PARTIAL.  
 KW Oxidoreductase; Peroxidase.  
 FT NON\_TER 20  
 SO SEQUENCE 20 AA: 2193 MW: 08178FCD2782E765 CRC64;

Query Match 17.2%; Score 25; DB 1; Length 20;  
 Best Local Similarity 41.7%; Pred. No. 1e+03;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 8 DVSIAFSEISVG 19  
 DB 9 EVTLQKEVKVG 20

RESULT 4  
 ID RM12\_YEAST STANDARD; PRT; 15 AA.  
 AC P36522;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L12 (YML12) (FRAGMENT).  
 GN MRPL12.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=91285106; PubMed=2060626;  
 RA Grohmann L., Graack H.-R., Kruft V., Choll T., Goldschmidt-Reisin S.,  
 RA Kitakawa M.;  
 RT "Extended N-terminal sequencing of proteins of the large ribosomal  
 RT subunit from yeast mitochondria.";  
 RL FEBS Lett. 284:51-56(1991).  
 DR PIR: S17261; S17261.  
 DR SGD: L0002687; MRPL12.  
 KW Ribosomal protein; Mitochondrion.  
 FT NON\_TER 15  
 SO SEQUENCE 15 AA: 1851 MW: 74BCD9FEDDB3900 CRC64;

Query Match 16.6%; Score 24; DB 1; Length 15;  
 Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 GYHLDVSLAFS 14  
 DB 4 GYFEVYIVRFN 14

RESULT 5  
 ID XYLA\_STRVN STANDARD; PRT; 12 AA.  
 AC P14405;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE XYLOSE ISOMERASE (EC 5.3.1.5) (FRAGMENT).  
 OS Streptomyces violaceoruber.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=1935;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=LMG 7183;  
 RX MEDLINE=90104230; PubMed=2604694;  
 RA Vangryperre W., Ampe C., Kersters-Hinderson H., Tempst P.;  
 RT "Single active-site histidine in D-xylose isomerase from *Streptomyces*  
 RT *violaceoruber*. Identification by chemical derivatization and peptide  
 RT mapping.";  
 RL Blochem. J. 263:195-199(1989).  
 CC -1- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.  
 CC -1- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.  
 DR HSSP: P37031; IDXI.  
 DR InterPro: IPR001998; Xylose\_isom.  
 DR PROSITE: PS00172; XYLOSE\_ISOMERASE\_1; PARTIAL.

DR PROSITE: PS00173; XYLOSE-ISOMERASE\_2, PARTIAL.  
 KW Isomerases; Pentose shunt; Xylose metabolism; Magnesium.  
 FT NON\_TER 1 1  
 FT ACT\_SITE 5 5  
 FT NON\_TER 12 12  
 FT SEQUENCE 12 AA; 1375 MW; E749268EB1AAAA1 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 12;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 19 GAFFNKD 26  
 Db 1 GVFPHDD 8

RESULT 6  
 BP37\_LEUMA STANDARD; PRT; 13 AA.  
 AC P81754;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE BRAIN PROTEIN 37F3.  
 OS Leucophaea maderae (Madelira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidae; Blaberidae; Leucophaea.  
 NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Brain;  
 RX MEDLINE=97269266; PubMed-9114447;  
 RA Muren J.E., Naessel D.R.;  
 RT "Seven tachykinin-related peptides isolated from the brain of the  
 RT madelira cockroach: evidence for tissue-specific expression of  
 RT isoforms.";  
 RL Peptides 18:7-15(1997).  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- MASS SPECTROMETRY: MW=1435.7; METHOD=MALDI.  
 SQ SEQUENCE 13 AA; 1436 MW; 8E532C9DE9A7D1B5 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 13;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 13 FSEISVGA 21  
 Db 3 FEESTVSAE 11

RESULT 7  
 AGI\_EUPCH STANDARD; PRT; 18 AA.  
 AC P33888;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DE LECTIN (FRAGMENT).  
 OS Euphorbia characias (spurge).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.  
 NCBI\_TaxID=3991;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=latex;  
 RX MEDLINE=93357266; PubMed-8353129;  
 RA Stierpe F., Licastro F., Morini M.C., Parente A., Savino G.,  
 RA Abondanza A., Bolognesi A., Falasca A.I., Rossi C.A.;  
 RT "Purification and partial characterization of a mitogenic lectin from

RT the latex of Euphorbia marginata.";  
 RL Biochim. Biophys. Acta 1158:33-39(1993).  
 CC -1- FUNCTION: LECTIN THAT BINDS GALACTOSE, GALACTOSE-CONTAINING SUGARS  
 CC AND GENTIOBIOSIDE. IT IS STRONGLY MITOGENIC FOR HUMAN T LYMPHOCYTES.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- PTM: N-GLYCOSYLATED.  
 CC -1- SIMILARITY: TO E.MARGINATA LECTIN.  
 DR PIR; S36121; S36121.  
 KW Lectin.  
 FT NON\_TER 18 18  
 FT SEQUENCE 18 AA; 1923 MW; C6F6A1A7B2AB124F CRC64;

Query Match 15.9%; Score 23; DB 1; Length 18;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GYHLDV 9  
 Db 12 GYXVDV 17

RESULT 8  
 ANFC\_CHICK STANDARD; PRT; 22 AA.  
 ID ANFC\_CHICK  
 AC P21805;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE C-TYPE NATRIURETIC PEPTIDE (CNP).  
 GN NPFC.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=9113166; PubMed-1989595;  
 RA Arimura J.J., Minamino N., Kangawa K., Matsuo H.;  
 RT "Isolation and identification of C-type natriuretic peptide in  
 RT chicken brain.";  
 RL Biochem. Biophys. Res. Commun. 174:142-148(1991).  
 CC -1- FUNCTION: VASORELAXANT ACTIVITY. HAS A CGMP-STIMULATING ACTIVITY.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.  
 DR PIR; JTO581; JTO581.  
 DR InterPro; IPR000663; Natr\_peptide.  
 DR Pfam; PF00212; ANP; 1.  
 DR PRINTS; PR00710; NATPEPTIDES.  
 DR SMART; SM00183; NAT\_PEP; 1.  
 DR PROSITE; PS00263; NATRIURETIC\_PEPTIDE; 1.  
 KW Vasoactive.  
 FT DISULFID  
 FT SEQUENCE 22 AA; 2244 MW; FC2A4706D8DAC025 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 22;  
 Best Local Similarity 57.1%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CLGYHLD 8  
 Db 6 CFGVKLD 12

RESULT 9  
 FMK7\_PSEAE STANDARD; PRT; 23 AA.  
 ID FMK7\_PSEAE  
 AC Q53391;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

```

DE 15-JUL-1998 (Rel. 36, last annotation update)
DE FTMBRAL PROTEIN (PILIN) (STRAIN KB7) (FRAGMENT).
GN PIL.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KB7;
RA MEDLINE=94103636; PubMed=7903973;
RX IPATSG;
RT "A multicenter comparison of methods for typing strains of
RT Pseudomonas aeruginosa predominantly from patients with cystic
RT fibrosis. The International Pseudomonas aeruginosa Typing Study
RT Group."
RL J. Infect. Dis. 169:134-142(1994).
RN [2]
RP STRUCTURE BY NMR OF 7-23.
RX STRAIN=KB7;
RX MEDLINE=96110702; PubMed=8845350;
RA Campbell A.P., McInnes C., Hodges R.S., Sykes B.D.;
RT "Comparison of NMR solution structures of the receptor binding
RT domains of Pseudomonas aeruginosa pilI strains PAO, KB7, and PAK:
RT implications for receptor binding and synthetic vaccine design."
RL Biochemistry 34:16255-16268(1995).
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; S67809; CAB3861.1; -
DR PDB; 1KB7; 29-JAN-96.
DR PDB; 1KB8; 29-JAN-96.
DR InterPro: IPR001082; Pilin.
DR InterPro: IPR001120; Prok_N_methyltn.
DR Pfam: PF00114; pilin; 1.
DR PROSITE: PS00409; PROKAR_NTER_METHYL; PARTIAL.
KW fimbria; 3D-structure.
FT NON_TER
FT DISULFID
FT SEQUENCE 23 AA; 2415 MW; DOBE77514AF041CF CRC64;

Query Match
Best Local Similarity 15.9%; Score 23; DB 1; Length 23;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 17 SVGAENKDDC 27
DB 11 TVDAKFRPNC 21

RESULT 10
COXC_THUOB
ID COXC_THUOB STANDARD; PRT; 24 AA.
AC P80973;
DR 01-NOV-1997 (Rel. 35, Created)
DR 01-NOV-1997 (Rel. 35, last sequence update)
DR 30-MAY-2000 (Rel. 39, last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-2 (EC 1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

```

OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeltch F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver."
RL Eur. J. Biochem. 248:99-103(1997).
CC -1- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
FT NON_TER
FT SEQUENCE 24 AA; 2903 MW; 20998FB91F22E43B CRC64;

Query Match
Best Local Similarity 15.9%; Score 23; DB 1; Length 24;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 22 FNKDD 26
DB 19 FNKSD 23

RESULT 11
POQA_PSEFL
ID POQA_PSEFL STANDARD; PRT; 24 AA.
AC P55171;
DR 01-OCT-1996 (Rel. 34, Created)
DR 01-OCT-1996 (Rel. 34, last sequence update)
DR 01-OCT-1996 (Rel. 34, last annotation update)
DE COENZYME PQQ SYNTHESIS PROTEIN A.
GN POQA.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHAO;
RX MEDLINE=96064397; PubMed=8526497;
RA Schneider U., Keel C., Defago G., Haas D.;
RT "Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHAO;
RT mutational inactivation of the genes results in overproduction of the
RT antibiotic pyoverdinin."
RL Appl. Environ. Microbiol. 61:3856-3864(1995).
CC -1- FUNCTION: REQUIRED FOR COENZYME PYROLO-QUINOLINE-QUINONE (PQQ)
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
CC ARE NECESSARY FOR THE SYNTHESIS.
CC -1- SIMILARITY: TO OTHER BACTERIAL PQA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; X87299; CAA60732.1; -
DR EMBL; X87299; CAA60731.1; ALT_INIT.
DR PQQ.
FT SEQUENCE 24 AA; 2871 MW; ED13AB79B2EA3EBE CRC64;

```



05 *Klebsiella pneumoniae*.  
0C Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae

```

OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCIC 418;
RA MEDLINE=92212283; PubMed=1313537;
RT Meulenber J.J.M., Sellink E., Riegan N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqr
operon."
RL Mol. Gen. Genet. 232:284-294(1992).
CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
CC ARE NECESSARY FOR THE SYNTHESIS.
CC -1- SIMILARITY: TO OTHER BACTERIAL PQA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58778; CAA41579.1; .
DR PIR; S20453; S20453.
KW PQQ.
SQ SEQUENCE 23 AA; 2764 MW; ACGB321460871C5D CRC64;

Query Match 15.2%; Score 22; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 3.5e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 12 AFSEISVGA 21
DB 6 AFIDRLGLE 15

RESULT 16
PEPM_SKECO
ID PEPM_SKECO STANDARD; PRT; 25 AA.
AC 096810;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE CYTOCHROME B6-F COMPLEX SUBUNIT VII (CYTOCHROME B6F COMPLEX SUBUNIT
DE PEPM) (FRAGMENT).
GN PEPM OR YCF31.
OS Skeletonema costatum.
OC Chloptoplast.
OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Thalassiosirophyceae; Thalassiosirales; Skeletonemataceae;
OC Skeletonema.
OX NCBI_TaxID=2843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIES-323;
RA Tada N., Otsuka S., Oyaizu H., Matsumoto S.;
RT "Plastid DNA sequences of Skeletonema costatum NIES 323."
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CYTOCHROME B6-F COMPLEX SUBUNIT.
CC -1- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE-ASSOCIATED (BY
CC SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE PEPM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AJ132266; CAA10631.1; .
KW Electron transport; Chloroplast; Respiratory chain; Thylakoid;
KW Transmembrane.
FT NON-TER 1 1
FT TRANSMEM 2 22
SQ SEQUENCE 25 AA; 2630 MW; 99ECF1B6CA3113AA CRC64;

Query Match 15.2%; Score 22; DB 1; Length 25;
Best Local Similarity 38.5%; Pred. No. 3.8e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 9 VSLAFSEISVGA 21
DB 13 LSLGFAALLVQGE 25

RESULT 17
CX7A_CONTE
ID CX7A_CONTE STANDARD; PRT; 27 AA.
AC P24160;
DT 01-MAR-1992 (Rel. 21, Created)
DT 30-MAY-2000 (Rel. 38, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE CONOTOXIN TXVIIA (TXIIA).
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE.
RC STRAIN-NEOVITCARIUS; TISSUE-Venom;
RX MEDLINE=92104183; PubMed=1761058;
RA Fainzilber M., Gordon D., Hesson A., Spira M.E., Zlotkin E.;
RT "Mollusc-specific toxins from the venom of Conus textile
RT neovicarius."
RL Eur. J. Biochem. 202:589-595(1991).
RN [2]
RP REVISIONS TO 1 AND C-TERMINAL, AND MASS-SPECTROMETRY.
RX MEDLINE=97022130; PubMed=8868490;
RA Nakamura T., Yu Z., Fainzilber M., Burlingame A.L.;
RT "Mass spectrometric-based revision of the structure of a cysteine-rich
RT peptide toxin with gamma-carboxyglutamic acid, TxVIIA, from the sea
RT snail, Conus textile."
RL Protein Sci. 5:524-530(1996).
CC -1- FUNCTION: POTENT NEUROTOXIN. MAY EXERT ITS EFFECTS AT THE LEVEL
CC OF THE NEUROMUSCULAR JUNCTION.
CC -1- P.TM: THREE DISULFIDE BONDS ARE PRESENT.
CC -1- MASS SPECTROMETRY: MW=3088.9; METHOD-ELECTROSPRAY.
DR PIR; S19620; S19620.
KW Venom; Neurotoxin; Vitamin K; Gamma-carboxyglutamic acid; Amidation.
FT MOD_RES 9 9 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 13 13 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 27 27 AMIDATION.
SQ SEQUENCE 27 AA; 3008 MW; D7A49781300FE1E7 CRC64;

Query Match 14.8%; Score 21.5; DB 1; Length 27;
Best Local Similarity 26.9%; Pred. No. 4.8e+03;
Matches 7; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

OY 2 CIGYHLDVSLAFSEISVGAEPKDDC 27
DB 1 CGGYS-----TYCEVDSECCSDNC 19

RESULT 18
TKL4_LOCM1
ID TKL4_LOCM1 STANDARD; PRT; 10 AA.
AC P30250;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)

```

DT 01-NOV-1995 (Rel. 32, last annotation update)  
 DE LOCUSTATACHYKININ IV (TK-IV).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthoptera; Orthopteroidea; Orthoptera; Caelifera.  
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.  
 OX NCBI\_TaxId=7004;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-Brain;  
 RX MEDLINE=91219696; PubMed=2132575;  
 RA Schoofs L., Holman G.M., Hayes T.K., Kochansky J.P., Nachman R.J.,  
 de Loof A.;  
 RT "Locusttachykinin III and IV: two additional insect neuropeptides  
 with homology to peptides of the vertebrate tachykinin family.";  
 RL Regul. Pept. 31:199-212(1990).  
 CC -1- FUNCTION: MYOACTIVE PEPTIDE. STIMULATES THE CONTRACTION OF THE  
 OVIDUCT AND FOREGUT.  
 CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
 CC PIR: B60073; ECIQ4M.  
 DR Tachykinin; Neuropeptide; Amladation.  
 KW MOD.RES 10 10  
 FT SEQUENCE 10 AA; 1040 MW; 9E52CD71E9C87735 CRC64;

Query Match 14.5%; Score 21; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGYH 6  
 1111  
 Db 4 LGFH 7

RESULT 19  
 ID FIBA\_SHEEP STANDARD; PRT; 19 AA.  
 AC P14451;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, last sequence update)  
 DT 01-MAY-1991 (Rel. 18, last annotation update)  
 DE FIBRINOPEPTIDE A.  
 OS Ovis aries (Sheep), and Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxId=9940, 9923;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals.";  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR HSSP: P02671; 1BBR.  
 KM Blood coagulation; Plasma.  
 FT NON\_TER 19 19  
 SO SEQUENCE 19 AA; 1848 MW; 99C98873D1B059C5 CRC64;

Query Match 14.5%; Score 21; DB 1; Length 19;  
 Best Local Similarity 80.0%; Pred. No. 4.1e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 18 VGAEF 22  
 111111

Db 7 VGGEF 11

RESULT 20  
 ID V101\_VACCV STANDARD; PRT; 20 AA.  
 AC P16714;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, last sequence update)  
 DT 01-FEB-1991 (Rel. 17, last annotation update)  
 DE PROTEIN I1 (FRAGMENT).  
 GN I1L.  
 OS Vaccinia virus (strain WR).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 OX NCBI\_TaxId=10254;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88215015; PubMed=2835495;  
 RA Schmitt J.F.C., Stuenkelberg H.G.;  
 RT "Sequence and transcriptional analysis of the vaccinia virus HindIII  
 I fragment";  
 RL J. Virol. 62:1889-1897(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J03399; AAB59803.1; -.  
 DR PIR: A29889; A29889.  
 KW Late protein.  
 RN NON\_TER 20 20  
 FT SEQUENCE 20 AA; 2241 MW; 6E97AA78D505C4B1 CRC64;

Query Match 14.5%; Score 21; DB 1; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 4.3e+03;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 11 LAFSEISVGA 20  
 111111  
 Db 8 LVFNSISARA 17

RESULT 21  
 ID RL10\_KLEPN STANDARD; PRT; 22 AA.  
 AC P41190;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, last sequence update)  
 DT 20-AUG-2001 (Rel. 40, last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L10 (FRAGMENT).  
 GN RPLJ.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 OX NCBI\_TaxId=573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 5054;  
 RA Znyvoloup A.N.;  
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X74445; CAA52456.1; -.
DR PIR: S35975; S35975.
DR PIR: S35976; S35976.
DR InterPro: IPR002363; Ribosomal_L10_eub.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT_MET 0
FT NON_TER 22
SQ SEQUENCE 22 AA; 2268 MW; 113D96520A02CD78 CRC64;

Query Match
Best Local Similarity 14.5%; Score 21; DB 1; Length 22;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 SEISVGA 20
DB 15 SEVAKGA 21

RESULT 22
RL10_ENTCL STANDARD; PRT; 23 AA.
AC Q47608; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L10 (FRAGMENT).
GN RPLJ.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhyvoloup A.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X74444; CAA52455.1; -.
DR InterPro: IPR002363; Ribosomal_L10_eub.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT_MET 0
FT NON_TER 23
SQ SEQUENCE 23 AA; 2355 MW; 39413D96520A02CD CRC64;

Query Match
Best Local Similarity 14.5%; Score 21; DB 1; Length 23;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 SEISVGA 20
DB 15 SEVAKGA 21

RESULT 23
CAMT_PINPS STANDARD; PRT; 24 AA.
ID CAMT_PINPS
AC P81081;
DT 15-JUL-1998 (Rel. 36, Created)

```

```

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE CAFFEYL-COA O-METHYLTRANSFERASE (EC 2.1.1.104) (TRANS-
DE CAFFEYL-COA 3-O-METHYLTRANSFERASE) (CCOAMT) (CCOAMOT) (WATER STRESS
DE RESPONSIVE PROTEIN 13) (FRAGMENTS).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=98418576; PubMed=9747804;
RA Costa P., Bahman N., Frigerio J.-M., Kremer A., Plomion C.;
RT "Water-deficit-responsive proteins in maritime pine."
RN Plant Mol. Biol. 38:587-596(1998).
RN [2]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins."
RN Electrophoresis 20:1098-1108(1999).
CC -!- FUNCTION: INVOLVED IN THE REINFORCEMENT OF THE PLANT CELL WALL
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + CAFFEYL-COA =
CC S-ADENOSYL-L-HOMOCYSTEINE + FERULOYL-COA.
CC -!- PATHWAY: LIGNIN BIOSYNTHESIS.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC IS: 5.2. ITS MW IS: 30 KDA.
KW Lignin biosynthesis; Transferase; Methyltransferase.
FT NON_TER 1
FT NON_CONS 9
FT NON_TER 24
SQ SEQUENCE 24 AA; 2447 MW; AAB8042F9EB0CF0 CRC64;

Query Match
Best Local Similarity 14.5%; Score 21; DB 1; Length 24;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 SLAFSEISVG 19
DB 9 NIEISQIPVG 18

RESULT 24
ANF_ANGJA STANDARD; PRT; 27 AA.
ID ANF_ANGJA
AC P18144;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATRIAL NATRIURETIC FACTOR (ANF) (ATRIAL NATRIURETIC PEPTIDE) (ANP).
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguillidae; Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart atrium;
RX MEDLINE=90026430; PubMed=2529857;
RA Takei Y., Takahashi A., Watanabe T.X., Nakajima K., Sakakibara S.;
RT "Amino acid sequence and relative biological activity of eel atrial
RL natriuretic peptide."
CC Biochem. Biophys. Res. Commun. 164:537-543(1989).
CC -!- FUNCTION: EXHIBITS NATRIURETIC AND VASODEPRESSOR ACTIVITY. HAS A
CC CGMP-STIMULATING ACTIVITY.

```

CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.  
 DR PIR; A3431; A3431.  
 DR InterPro: IPR000663; Natr\_peptide.  
 DR Pfam: PF00212; ANP; 1.  
 DR PRINTS: PR00710; NATPEPTIDES.  
 DR SMART; SM00183; NAT\_PEP; 1.  
 DR PROSITE; PS00263; NATRIURETIC\_PEPTIDE; 1.  
 DR Vasoactive.  
 KW DISULFID  
 FT DISULFID 7 23  
 SQ SEQUENCE 27 AA; 2792 MW; 9EF510CCF282B4D CRC64;

Query Match 14.5%; Score 21; DB 1; Length 27;  
 Best Local Similarity 57.1%; Pred. No. 5.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CLGYHLD 8  
 Db 7 CFSGKLD 13

## RESULT 25

CD59\_SHEEP  
 ID CD59\_SHEEP STANDARD; PRT; 27 AA.  
 AC P58020;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CD59 GLYCOPROTEIN (MEMBRANE ATTACK COMPLEX INHIBITION FACTOR) (MACIF)  
 DE (MAC-INHIBITORY PROTEIN) (MAC-IP) (PROTECTIN) (COMPLEMENT INHIBITORY  
 DE PROTEIN) (CIP) (FRAGMENT).  
 GN CD59.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Erythrocyte;  
 RX MEDLINE=93239201; Pubmed=7682985;  
 RA van den Berg C.W., Harrison R.A., Morgan B.P.;  
 RT "The sheep analogue of human CD59: purification and characterization  
 of its complement inhibitory activity.";  
 RL Immunology 78:349-357(1993).  
 CC -1- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK  
 CC COMPLEX (MAC) ACTION. ACTS BY BINDING, IN A SPECIES SPECIFIC  
 CC MANNER, TO THE C8 AND/OR C9 COMPLEMENTS OF THE ASSEMBLING MAC,  
 CC THEREBY PREVENTING INCORPORATION OF THE MULTIPLE COPIES OF C9  
 CC REQUIRED FOR COMPLETE FORMATION OF THE OSMOTIC PORE.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ERYTHROCYTES AND LYMPHOCYTES.  
 CC NOT DETECTED IN PLATELETS.  
 DR PROSITE; PS00983; LY6\_UPAR; PARTIAL.  
 KW Antigen; Glycoprotein.  
 FT NON\_TER 1  
 FT DISULFID 3 25  
 FT DISULFID 6 12  
 FT DISULFID 17 17  
 FT CARBOHYD 3 3  
 FT UNSURE 6 6  
 FT UNSURE 17 17  
 FT UNSURE 18 18  
 FT UNSURE 25 25  
 FT NON\_TER 27  
 SQ SEQUENCE 27 AA; 3025 MW; 92E6C978EC7663FE CRC64;

Query Match 14.5%; Score 21; DB 1; Length 27;  
 Best Local Similarity 30.0%; Pred. No. 5.8e+03;  
 Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 18 VGAEPKDDC 27  
 Db 16 INCTXNODAC 25

Search completed: January 6, 2002, 09:17:45  
 Job time: 278 sec

**THIS PAGE BLANK (USPTO)**



OC Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;  
 OC Oxytrichidae; Oxytricha.  
 OX NCBI\_TaxID=5947;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95345843; PubMed=7620461;  
 RA Irato P., Piccini E., James P., Ammermann D.;  
 RT "Evidence of a cadmium-thionein and the glycine cleavage system in  
 RL Oxytricha granulifera.";  
 SO J. Eukaryot. Microbiol. 42:376-378(1995).  
 SEQUENCE 19 AA; 2381 MW; 4C0B562B50A0984 CRC64;

Query Match  
 Best Local Similarity 21.4%; Score 31; DB 5; Length 19;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 OY 13 FSEISVGAEPNKD 25  
 Db 5 FTEDAEWVEYKND 17

RESULT 3  
 ID 094182 PRELIMINARY; PRT; 22 AA.  
 AC 094182;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE COSMID C47C12.  
 GN C47C12.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Keshaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lighting J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton M., Smith C., Sonnenhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Connell M.;  
 RN [3]  
 RP Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RT Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U67951; AAB07573.1; -.  
 SO SEQUENCE 22 AA; 2535 MW; 35EA30AF96C960B5 CRC64;

Query Match  
 Best Local Similarity 21.4%; Score 31; DB 5; Length 22;  
 Matches 9; Conservative 2; Mismatches 5; Indels 2; Gaps 1;  
 OY 6 HLDV-SLAFSEISVGA 21  
 Db 4 HSDIGNSLOFSRISLKN 21

RESULT 4  
 ID 041052 PRELIMINARY; PRT; 26 AA.  
 AC 041052;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE VAR. AMPLIFOLIA DNA SEQUENCE WITH PARTIAL CDS (FRAGMENT).  
 OS Porphyra spiralis.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
 OX NCBI\_TaxID=31350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PSA-B/UBAT0BA;  
 RX MEDLINE=94224117; PubMed=8170361;  
 RA Oliveira M.C., Ragan M.A.;  
 RT "Variant forms of a group I intron in nuclear small-subunit rRNA genes  
 RL of the marine red alga Porphyra spiralis var. amplifolia.";  
 DR EMBL; L26175; AAA72417.1; -.  
 FT NON\_TER  
 SO SEQUENCE 26 AA; 2821 MW; 9F53369E5A2254F CRC64;

Query Match  
 Best Local Similarity 20.7%; Score 30; DB 10; Length 26;  
 Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 OY 5 YHLDVSLAFSEISVG 19  
 Db 5 YHLEGEVYTRFSVG 19

RESULT 5  
 ID 09R569 PRELIMINARY; PRT; 15 AA.  
 AC 09R569;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE HYDROXYLAMINE OXIDOREDUCTASE (FRAGMENT).  
 OS Nitrosomonas europaea.  
 OC Bacteria; Proteobacteria; beta subdivision;  
 OC Ammonia-oxidizing bacteria; Nitrosomonas.  
 OX NCBI\_TaxID=915;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93315429; PubMed=8325841;  
 RA Arciero D.M., Hooper A.B.;  
 RT "Hydroxylamine oxidoreductase from Nitrosomonas europaea is a multimer  
 RT of an octa-heme subunit.";  
 RL J. Biol. Chem. 268:14645-14654(1993).  
 DR HSSP; Q50923; 1F9J.  
 DR InterPro; IPR000345; Cytochrome\_c.  
 DR PROSITE; PS00190; CYTOCHROME\_C.  
 SO SEQUENCE 15 AA; 1687 MW; 983D4B8A1369849 CRC64;

Query Match  
 Best Local Similarity 19.3%; Score 28; DB 2; Length 15;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 OY 2 CIGYHLDVS 10  
 Db 5 CIDCHVDVN 13

RESULT 6  
 ID 09TWK8 PRELIMINARY; PRT; 19 AA.  
 AC 09TWK8;



```
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE HAPTOGLOBIN-RELATED PROTEIN BETA SUBUNIT (FRAGMENT).
OC Trypanosoma brucei.
OC Eukaryota; Euzoosoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
ON NCBI_TaxID=5691;
RN [1]
RP MEDLINE=95232503; Pubmed=7716520;
RA Smith A.B., Esko J.D., Hajduk S.L.;
RT "Killing of trypanosomes by the human haptoglobin-related protein.";
RL Science 268:284-286(1995).
SQ SEQUENCE 19 AA; 2055 MW; 65BD135667C94056 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 5; Length 19;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LGYHLDSVLAFL 13
Db 2 LGHLDKSGSF 12

RESULT 7
Q82392 PRELIMINARY; PRT; 21 AA.
AC 082392;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
ON NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eiraku N., Novoa P., Monken C., Oliveira M.D., Oliveira O.D.,
RA Ishak R., Oliveira M.P., Laureiro P., Ishak M., Acevedo V.,
RA Hammershlag N., Zhu S.W., Kubo T., Hall W.W.;
RT "Identification and characterization of a new and distinct molecular
RT subtype of human T-cell lymphotropic virus type 2.";
RL J. Virol. 70:1481-1492(1996).
DR EMBL: U32871; AAB04905.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2383 MW; 11F39408575CF6D0 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 12; Length 21;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 13 FSEISVGAERFKDD 26
Db 2 YTNIPVSIIFNKEE 15

RESULT 8
Q82410 PRELIMINARY; PRT; 21 AA.
AC 082410;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
ON NCBI_TaxID=11909;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eiraku N., Novoa P., Monken C., Oliveira M.D., Oliveira O.D.,
RA Ishak R., Oliveira M.P., Laureiro P., Ishak M., Acevedo V.,
RA Hammershlag N., Zhu S.W., Kubo T., Hall W.W.;
RT "Identification and characterization of a new and distinct molecular
RT subtype of human T-cell lymphotropic virus type 2.";
RL J. Virol. 70:1481-1492(1996).
DR EMBL: U32883; AAB04923.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2383 MW; 11F39408575CF6D0 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 12; Length 21;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 13 FSEISVGAERFKDD 26
Db 2 YTNIPVSIIFNKEE 15

RESULT 9
Q82411 PRELIMINARY; PRT; 21 AA.
AC 082411;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
ON NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eiraku N., Novoa P., Monken C., Oliveira M.D., Oliveira O.D.,
RA Ishak R., Oliveira M.P., Laureiro P., Ishak M., Acevedo V.,
RA Hammershlag N., Zhu S.W., Kubo T., Hall W.W.;
RT "Identification and characterization of a new and distinct molecular
RT subtype of human T-cell lymphotropic virus type 2.";
RL J. Virol. 70:1481-1492(1996).
DR EMBL: U32884; AAB04924.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2383 MW; 11F39408575CF6D0 CRC64;

Query Match
Best Local Similarity 19.3%; Score 28; DB 12; Length 21;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 13 FSEISVGAERFKDD 26
Db 2 YTNIPVSIIFNKEE 15

RESULT 10
Q82412 PRELIMINARY; PRT; 21 AA.
AC 082412;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE T-CELL LYMPHOTROPIC VIRUS TYPE 2 TAX (FRAGMENT).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
ON NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190544; Pubmed=8627666;
RA Eiraku N., Novoa P., Monken C., Oliveira M.D., Oliveira O.D.,
```

KLIFFFEN M., de Jong P.T., Luiders T.M.,  
"Protein analysis of human maculae in relation to age-related  
maculopathy",  
Lab. Invest. 73:267-272(1995).

3 LGYHLDVSL-AFSEISVGAEFNKD 25

Db 1 MSFRINTNIALTSHAVGVONNRD 24

## RESULT 15

09DE23 PRELIMINARY; PRT; 20 AA.  
AC 09DE23;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE UORF2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=20528616; PubMed=11073974;  
RX Kobayashi M., Yu R.T., Yasuda K., Umesono K.;  
RT "Cell-type-specific regulation of the retinoic acid receptor mediated  
RT by the orphan nuclear receptor TLX";  
RL Mol. Cell. Biol. 20:8731-8739(2000).  
DR EMBL; AF220160; AAG35363.1; -;  
DR InterPro; IPR000345; Cytochrome\_bind.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
SO SEQUENCE 20 AA; 2261 MW; AA7738B0BE3C482 CRC64;

Query Match 17.9%; Score 26; DB 13; Length 20;  
Best Local Similarity 66.7%; Pred. No. 3.e+03;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 6 HLDVSLAFSEISVG 14  
Db 5 HLDQQLAFS 13

## RESULT 16

063076 PRELIMINARY; PRT; 21 AA.  
AC 063076;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE RAT ALPHA-2 TYPE I COLLAGEN (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=85122694; PubMed=6395893;  
RX Genovese C., Rowe D., Kream B.;  
RT "Construction of DNA sequences complementary to rat alpha 1 and alpha  
RT 2 collagen mRNA and their use in studying the regulation of type I  
RT collagen synthesis by 1,25-dihydroxyvitamin D.";  
RL Biochemistry 23:6210-6216(1984).  
DR EMBL; M12200; AAA40835.1; -;  
DR InterPro; IPR000885; FIB\_collagen\_C.  
DR Pfam; PF01410; COLFI; 1.  
FT NON\_TER 1 21  
FT NON\_TER 1 21  
SO SEQUENCE 21 AA; 2358 MW; B86A6ECCEB46D60 CRC64;

Query Match 17.9%; Score 26; DB 11; Length 21;  
Best Local Similarity 33.3%; Pred. No. 3.e+03;  
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 5 YHLDVSLAFSEISVG 19  
Db 7 YHCKNSIAYLDEETG 21

## RESULT 17

09Z274 PRELIMINARY; PRT; 22 AA.  
AC 09Z274;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE NADH1 (FRAGMENT).  
OS Phlebotomus bergeroti.  
OS Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Psychodidae;  
OC Psychodidae; Phlebotomus.  
OX NCBI\_TaxID=59273;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Essegghir E., Ready P.D., Killick-Kendrick R., Ben Ismail R.;  
RT "Mitochondrial haplotypes and phylogeography of Phlebotomus vectors of  
RT Leishmania major";  
RL Insect Mol. Biol. 6:0-0(1997).  
DR EMBL; U94474; AAD09262.1; -;  
DR Mitochondrion.  
KW Mitochondrion.  
FT NON\_TER 1 1  
SO SEQUENCE 22 AA; 2453 MW; CD5B7EDC1A37035A CRC64;

Query Match 17.9%; Score 26; DB 8; Length 22;  
Best Local Similarity 35.7%; Pred. No. 3.e+03;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 6 HLDVSLAFSEISVG 19  
Db 3 YLPASLCFISLFTG 16

## RESULT 18

09Z272 PRELIMINARY; PRT; 22 AA.  
AC 09Z272;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE NADH1 (FRAGMENT).  
OS Phlebotomus papatasi.  
OS Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Psychodidae;  
OC Psychodidae; Phlebotomus.  
OX NCBI\_TaxID=29031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Essegghir E., Ready P.D., Killick-Kendrick R., Ben Ismail R.;  
RT "Mitochondrial haplotypes and phylogeography of Phlebotomus vectors of  
RT Leishmania major";  
RL Insect Mol. Biol. 6:0-0(1997).  
DR EMBL; U94475; AAD09264.1; -;  
DR Mitochondrion.  
KW Mitochondrion.  
FT NON\_TER 1 1  
FT NON\_TER 1 1  
SO SEQUENCE 22 AA; 2453 MW; CD5B7EDC1A37035A CRC64;

Query Match 17.9%; Score 26; DB 8; Length 22;  
Best Local Similarity 35.7%; Pred. No. 3.e+03;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 6 HLDVSLAFSEISVG 19  
Db 3 YLPASLCFISLFTG 16

## RESULT 19

Q90123  
ID Q90123 PRELIMINARY; PRT; 27 AA.  
AC Q90123;  
RT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2/NS1 (GP68)]  
DE (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
NCBI\_Taxid=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of hypervariable region 1 of the viral  
genome and the sensitivity of hepatitis C virus to interferon alpha  
therapy."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -1 SUBUNIT: HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E2/NS1 IS  
CONTAINED WITHIN THE GENOME POLYPROTEIN THAT CONTAINS: CAPSID  
PROTEIN C; THE ENVELOPE GLYCOPROTEIN E1; ENVELOPE GLYCOPROTEIN  
E2/NS1; NONSTRUCTURAL PROTEINS NS2, NS3, NS4A, NS4B, NS5A AND  
NS5B. THE VIRION OF THE VIRUS IS COVERED BY A LIPIDPROTEIN ENVELOPE  
THAT CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E (BY  
SIMILARITY).  
CC -1 SIMILARITY: TO HEPATITIS C VIRUS NON-STRUCTURAL PROTEIN E2/NS1  
FAMILY.  
DR EMBL; AF166754; AAD52412.1; -.  
DR InterPro; IPR002531; HCV\_NS1.  
DR Pfam; PF01560; HCV\_NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1 1  
FT 27 27  
SQ SEQUENCE 27 AA; 2778 MW; 9D46112E115CA85A CRC64;

Query Match 17.9%; Score 26; DB 12; Length 27;  
Best Local Similarity 53.3%; Pred. No. 4.1e+03;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 9 VSLAFSEISVGAEPN 23  
Db 13 VSKFTSLFVGAQON 27

RESULT 20  
Q92820  
ID Q92820 PRELIMINARY; PRT; 13 AA.  
AC Q92820;  
RT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
NCBI\_Taxid=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SIVSM87;  
RX MEDLINE=96171634; PubMed=8574147;  
RA Kraiselburd E.N., Torres J.V.;  
RT "Properties of virus-like particles produced by SIV-chronically  
infected human cell clones."  
RL Cell. Mol. Biol. 41:41-52(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SIVSM87;  
RX MEDLINE=98105578; PubMed=9449524;  
RA Kraiselburd E.N., Salaman A., Beltran M., Rivera M., Oliver J.,  
Kessler M., Knezevich M., Rodriguez A., Biliska M., Montefiori D.,

RA Torres-Pauza L.J., Martinez I.;  
RT "Vaccine evaluation studies of replication-defective SIVsm87."  
RL Cell. Mol. Biol. 43:915-924(1997).  
DR EMBL; AF034690; AAC09235.1; -.  
KW Envelope protein.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1359 MW; 06E20308931E36D4 CRC64;

Query Match 17.2%; Score 25; DB 12; Length 13;  
Best Local Similarity 50.0%; Pred. No. 2.6e+03;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 CLGYHLDVSL 11  
Db 3 CLGNOLLNL 12

RESULT 21  
Q9R57  
ID Q9R57 PRELIMINARY; PRT; 16 AA.  
AC Q9R57;  
RT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE OUTER SHEATH PROTEIN.  
OS Treponema denticola.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
NCBI\_Taxid=158;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92041580; PubMed=1938897;  
RA Weinberg A., Holt S.C.;  
RT "Chemical and biological activities of a 64-kilodalton outer sheath  
protein from Treponema denticola strains."  
RL J. Bacteriol. 173:6935-6947(1991).  
SQ SEQUENCE 16 AA; 1640 MW; 3245DE6C11C5D124 CRC64;

Query Match 17.2%; Score 25; DB 2; Length 16;  
Best Local Similarity 35.7%; Pred. No. 3.3e+03;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 7 LDVSLAFSEISVGA 20  
Db 2 LDIALDGEAXINA 15

RESULT 22  
Q61946  
ID Q61946 PRELIMINARY; PRT; 24 AA.  
AC Q61946;  
RT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-NOV-1996 (TREMBLrel. 12, Last annotation update)  
DE NEURAL CELL ADHESION MOLECULE (FRAGMENT).  
GN NCAM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88283628; PubMed=3396534;  
RA Barbos J.A., Chai J.C., Steinmetz M., Goridis C.;  
RT "Distinct NCAM transcripts and alternative polyadenylation generates  
distinct NCAM transcripts and proteins in the mouse."  
RL EMBL; X07196; CAA30174.1; -.  
DR EMBL; MGI:97281; Ncam.  
FT NON\_TER 1 1  
SQ SEQUENCE 24 AA; 2457 MW; D529FAB2495D02D9 CRC64;



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:05:07 ; Search time 22.59 Seconds  
(without alignments)  
88.534 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFNKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 227461

Minimum DB seq length: 0  
Maximum DB seq length: 27

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_1101.\*  
1: /SID52/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 145   | 100.0       | 27     | 20    | AAV27328    |
| 2          | 35    | 24.1        | 15     | 21    | AAV77161    |
| 3          | 34    | 23.4        | 19     | 22    | AAE00613    |
| 4          | 34    | 23.4        | 20     | 18    | AAV35697    |
| 5          | 34    | 23.4        | 21     | 19    | AAV82186    |
| 6          | 34    | 23.4        | 21     | 22    | AAV673203   |
| 7          | 34    | 23.4        | 21     | 22    | AAV673204   |
| 8          | 34    | 23.4        | 26     | 21    | AAV96324    |
| 9          | 34    | 23.4        | 26     | 22    | AAV70374    |
| 10         | 34    | 23.4        | 27     | 21    | AAV37016    |
| 11         | 33    | 22.8        | 13     | 22    | AAU15328    |

|    |    |      |    |    |          |                     |
|----|----|------|----|----|----------|---------------------|
| 12 | 33 | 22.8 | 13 | 22 | AAU15338 | Schizophrenia-asso  |
| 13 | 33 | 22.8 | 15 | 19 | AAW38432 | Human dendritic ce  |
| 14 | 33 | 22.8 | 16 | 21 | AAW06316 | Human beta-amyloid  |
| 15 | 33 | 22.8 | 18 | 22 | AAE00609 | Beta-amyloid precu  |
| 16 | 33 | 22.8 | 19 | 22 | AAE00611 | Amyloid precursor   |
| 17 | 33 | 22.8 | 20 | 21 | AAV69714 | Beta-APP alpha-sec  |
| 18 | 33 | 22.8 | 20 | 21 | AAV69716 | Beta-APP alpha-sec  |
| 19 | 33 | 22.8 | 21 | 18 | AAW08361 | Beta-secretase sub  |
| 20 | 33 | 22.8 | 21 | 19 | AAW82188 | Fluorogenic protea  |
| 21 | 33 | 22.8 | 21 | 20 | AAV33755 | Synthetic oligopep  |
| 22 | 33 | 22.8 | 21 | 22 | AAV73205 | Protease binding s  |
| 23 | 33 | 22.8 | 21 | 22 | AAV73206 | Protease binding s  |
| 24 | 33 | 22.8 | 21 | 22 | AAV73316 | Protease indicator  |
| 25 | 33 | 22.8 | 21 | 22 | AAV47265 | Peptide 5-16'/SW, f |
| 26 | 32 | 22.1 | 10 | 20 | AAV47020 | Immunogenic peptid  |
| 27 | 32 | 22.1 | 10 | 22 | AAV43201 | Mycoplasma genital  |
| 28 | 32 | 22.1 | 10 | 22 | AAV43203 | Mycoplasma genital  |
| 29 | 32 | 22.1 | 10 | 22 | AAV43205 | Mycoplasma genital  |
| 30 | 32 | 22.1 | 11 | 21 | AAV54223 | HLA binding peptid  |
| 31 | 32 | 22.1 | 16 | 21 | AAW06315 | Human beta-amyloid  |
| 32 | 32 | 22.1 | 16 | 21 | AAW06317 | Human beta-amyloid  |
| 33 | 32 | 22.1 | 18 | 22 | AAE00608 | Beta-amyloid precu  |
| 34 | 32 | 22.1 | 20 | 16 | AAV64641 | HPF3 peptide deriv  |
| 35 | 32 | 22.1 | 20 | 21 | AAV69713 | Beta-APP alpha-sec  |
| 36 | 32 | 22.1 | 20 | 21 | AAV69715 | Beta-APP alpha-sec  |
| 37 | 32 | 22.1 | 20 | 22 | AAV55243 | Anti-HPV3 F1 regio  |
| 38 | 32 | 22.1 | 21 | 16 | AAV64642 | HPF3 peptide deriv  |
| 39 | 32 | 22.1 | 21 | 22 | AAV55244 | Anti-HPV3 F1 regio  |
| 40 | 32 | 22.1 | 22 | 16 | AAV64643 | HPF3 peptide deriv  |
| 41 | 32 | 22.1 | 22 | 22 | AAV55245 | Anti-HPV3 F1 regio  |
| 42 | 32 | 22.1 | 23 | 16 | AAV64644 | HPF3 peptide deriv  |
| 43 | 32 | 22.1 | 23 | 22 | AAV55246 | Anti-HPV3 F1 regio  |
| 44 | 32 | 22.1 | 24 | 16 | AAV64645 | HPF3 peptide deriv  |
| 45 | 32 | 22.1 | 24 | 22 | AAV55247 | Anti-HPV3 F1 regio  |

#### ALIGNMENTS

|          |   |
|----------|---|
| RESULT 1 |   |
| AAV27328 |   |
| ID       | AAV27328 standard; peptide: 27 AA.                                    |
| XX       |   |
| AC       | AAV27328;   |
| XX       |   |
| DT       | 05-NOV-1999 (first entry)   |
| XX       |   |
| DE       | Human C9 protein fragment (residues 359-384).                         |
| XX       |   |
| KW       | CD59 mediated complement; human; C9 protein; mimetic;                 |
| KW       | tumour therapy; complement-mediated inflammation; immune disorder;    |
| KW       | Immunovascularitis; rheumatoid arthritis; scleroderma; C5b-9 complex. |
| XX       |   |
| OS       | Homo sapiens.   |
| XX       |   |
| PN       | WO940115-A2.  |
| XX       |   |
| PD       | 12-AUG-1999.  |
| XX       |   |
| PF       | 09-FEB-1999; 99WO-US02802.  |
| XX       |   |
| PR       | 09-FEB-1998; 98US-0020393.  |
| XX       |   |
| PA       | (BLOO-) BLOOD CENT RES FOUND INC.                                     |
| PA       | (OKLA-) OKLAHOMA MEDICAL RES FOUND.                                   |
| XX       |   |
| PI       | Sims PJ;  |
| XX       |   |
| DR       | WPI; 1999-527301/44.  |
| XX       |   |
| PT       | Compounds modulating CD59 mediated complement activity, for           |
| PT       | treatment of, e.g. immunovascularitis                                 |
| XX       |   |





CC protein (APP) beta-secretase cleavage site (AAE0611) (NL/K1) mutant. APP  
 CC sequence containing this mutation is less susceptible to beta-secretase  
 CC cleavage. APP beta-secretase cleavage site is used to construct an  
 CC artificially engineered chimeric cassette comprising human caspase-3 with  
 CC interdomain linker replaced by Swedish mutant beta-secretase cleavage  
 CC site. This modified caspase-3 plays a pivotal role in Alzheimer's  
 CC disease. Caspases are a family of cysteine proteases, that participate  
 CC in the initiation and execution of apoptosis.  
 CC The present invention relates to a method for functional cloning of genes  
 CC encoding proteins or enzymes involved in proteolytic cleavage. The  
 CC invention is based on the use of caspase expression cassettes comprising  
 CC the coding sequence of a proteolytic cleavage site flanked by sequences  
 CC encoding two caspase subunits. A fusion polypeptide comprising a first  
 CC and a second caspase subunit, separated by a cleavage site not associated  
 CC in nature, is useful for cloning gene encoding enzymes involved in  
 CC proteolytic cleavage. An expression cassette containing fusion  
 CC polypeptide is used to identify a mutant cell line deficient in an  
 CC enzyme of interest and is also useful for diagnosis and suppression of  
 CC proliferation or metastases of a tumour cell characterised by  
 CC overexpression of a polypeptide (e.g. Cathepsin B or urokinase,  
 CC selectively expressed in the tumour cells). DNA encoding fusion  
 CC polypeptide is used in gene therapy.

XX Sequence 19 AA:

Query Match 23.4%; Score 34; DB 22; Length 19;

Best Local Similarity 50.0%; Pred. No. 65;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25

Db 7 sevkidaeefrhd 18

RESULT 4

AAW35697 AAW35697 standard; peptide; 20 AA.

XX AAW35697;

XX 13-MAY-1998 (first entry)

DE D. maculata antigen 5 peptide (residues 91-110).

XX Immunomodulatory peptide; vespid antigen 5; immunogenic; allergy;

KW vespid venom; white face hornet wasp; immunodominant peptide; T cell.

XX Dolichovespula maculata.

OS Synthetic.

XX WO9733910-A1.

XX 18-SEP-1997.

XX 11-MAR-1997; 97WO-US03753.

XX 11-MAR-1996; 96US-0614935.

XX (UYRO ) UNIV ROCKEFELLER.

XX King TP;

XX WPI; 1997-470817/43.

XX Vespid venom antigen 5 peptide fragments - useful to treat or

XX diagnose vespid venom sensitivity

XX Example 1; Figure 2; 73pp; English.

XX Sequences AAW35694-99 represent 15-20 residue peptides of white faced

XX hornet antigen 5 molecule. The invention relates to peptides derived

CC from vespid venom antigen 5 (VVS) that are antigenic for T cell

CC proliferation in mice immunised with VVS. The peptides can be used to  
 CC treat or diagnose vespid venom sensitivity e.g. to Dolichovespula  
 CC maculata (white face hornet), Vespula vulgaris (yellowjacket), V.  
 CC annulifrons (yellowjacket), D. arenaria (yellow hornet), Polistes  
 CC annularis (wasp), P. exclamans (wasp), V. crabro (European hornet), V.  
 CC flavopilosa (yellowjacket), V. germanica (yellowjacket), V. pennsylvanica  
 CC (yellowjacket), V. squamosa (yellowjacket), V. vidua (yellowjacket) and  
 CC P. fuscatus (paperwasp).

XX Sequence 20 AA:

Query Match 23.4%; Score 34; DB 18; Length 20;

Best Local Similarity 71.4%; Pred. No. 70;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 21 EFNKDDC 27

Db 5 dlnhdc 11

RESULT 5

AAW82186 AAW82186 standard; peptide; 21 AA.

XX AAW82186;

XX 18-FEB-1999 (first entry)

DE Fluorogenic protease indicator Swedish KM/NL AMLOID peptide #1.

XX Protease activity; fluorophore; detection; fluorogenic; cellular uptake;

KW conformation change.

XX Synthetic.

XX Key Location/Qualifiers

XX Modified-site 3 /label= Alb

XX /note= "alpha-aminoisobutyric acid, labelled as

XX Modified-site 4 /note= "epsilon-aminocaproic acid, labelled as

XX /note= "epsilon-aminocaproic acid, labelled as

XX Modified-site 16 /note= "epsilon-aminocaproic acid, labelled as

XX WO9837226-A1.

XX 27-AUG-1998.

XX 20-FEB-1998; 98WO-US03000.

XX 20-FEB-1997; 97US-0802981.

XX (ONCO-) ONCOIMMUNIN INC.

XX Komoriya A, Packard BS;

XX WPI; 1998-467579/40.

XX New fluorogenic compositions - containing 2 fluorophores separated

XX by a peptide comprising a protease binding site, used for detecting

XX protease activity in samples.

XX disclosure; Page 28; 90pp; English.

XX AAW82023-W82240 are peptides used in the construction of a fluorogenic

XX composition which is used for the detection of protease activity in

XX biological samples. The products can be used for the detection of

CC conformation changes in nucleic acids, oligosaccharides, polysaccharides, proteins, peptides, lipids, phospholipids, glycolipids,

CC glycoproteins, steroids or polymers. In addition, attachment of a  
 CC hydrophobic group to a molecule can be used to enhance uptake by cells.  
 CC The composition is composed of P = peptide comprising a protease binding  
 CC site for the protease, F1, F2 peptides = fluorophores where F1 is  
 CC attached to the amino terminal amino acid and F2 is attached to the  
 CC carboxyl terminal amino acid and S1, S2 peptides = when present, are  
 CC peptide spacers where S1, when present, is attached to the amino terminal  
 CC acid, and S2, when present, is attached to the carboxyl terminal amino  
 CC acid.  
 XX  
 SQ Sequence 21 AA;

Query Match 23.4%; Score 34; DB 19; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 74;  
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;  
 QY 14 SEISVGAEFN--KDD 26  
 ||::||| |||  
 Db 6 sevnldaefgcpkdd 20

## RESULT 6

AAAG73203 ID AAG73203 standard; Peptide: 21 AA.

AC AAG73203;

DT 14-AUG-2001 (first entry)

DE Protease binding site #137.

KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 KM thrombosis; haemophilia.

OS Synthetic.

FH Key Location/Qualifiers  
 FT Modified-site 3 /label= Alb

FT /note= "2-aminoisobutyric acid"

PN WO200118238-A1.

PD 15-MAR-2001.

PF 11-SEP-2000; 2000WO-US24882.

PR 10-SEP-1999; 99US-0394019.

PA (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

DR WPI; 2001-389573/41.

XX New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples -  
 XX  
 PS Disclosure; Page 27; 86pp; English.

XX The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore.  
 CC The peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.

SQ Sequence 21 AA;

Query Match 23.4%; Score 34; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 74;  
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;  
 QY 14 SEISVGAEFN--KDD 26  
 ||::||| |||  
 Db 6 sevnldaefgcpkdd 20

## RESULT 7

AAAG73204 ID AAG73204 standard; Peptide: 21 AA.

AC AAG73204;

DT 14-AUG-2001 (first entry)

DE Protease binding site #138.

KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 KM thrombosis; haemophilia.

OS Synthetic.

FH Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER

FT /note= "modified by fluorophore"

FT Modified-site 3 /label= Alb

FT /note= "2-aminoisobutyric acid"

PN WO200118238-A1.

PD 15-MAR-2001.

PF 11-SEP-2000; 2000WO-US24882.

PR 10-SEP-1999; 99US-0394019.

PA (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

DR WPI; 2001-389573/41.

XX New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples -  
 XX  
 PS Disclosure; Page 27; 86pp; English.

XX The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore.  
 CC The peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.

SQ Sequence 21 AA;

Query Match 23.4%; Score 34; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 74;  
 Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 14 SEISVGAERN--KDD 26  
 11::111111  
 Db 6 sevndaefgcpkdd 20

## RESULT 8

AA96324  
 ID AAY96324 standard; Peptide; 26 AA.

AC AAY96324;

DT 17-AUG-2000 (first entry)

DE Mammalian Bcl-2 homology domain 3 domain.

KM Mammal; apoptosis; cell death; BRC3; apoptosis promotion; Bcl-2  
 KM apoptosis inhibition; malignant cell; autoimmune disease.

OS Mammalia.

PN WO200026228-A1.

PD 11-MAY-2000.

PF 28-OCT-1999; 99WO-US25285.

PR 02-NOV-1998; 98US-0184168.

PA (CLON-) CLONTECH LAB INC.

PI Zhu L, Yin X, Chittenden T;

DR WPI: 2000-365560/31.

PT Novel polynucleotide encoding a BRC3 protein which is useful for  
 PT modulating apoptosis, especially in the treatment of cancer and  
 PT autoimmune diseases.

PS Disclosure; Fig 4; 47pp; English.

CC The present sequence is the mammalian Bcl-2 homology domain 3  
 CC (BH3) domain, which was used in a sequence alignment with the same  
 CC domain of a putative version of the mammalian apoptosis  
 CC regulator BRC3, which was designated BRC3-ORF2. The BRC3 protein,  
 CC nucleic acids and antibodies are suitable for use in promoting cell  
 CC death or for preventing apoptosis in malignant cells and those causing  
 CC autoimmune diseases.

SQ Sequence 26 AA;

Query Match 23 4%; Score 34; DB 21; Length 26;

Best Local Similarity 60.0%; Pred. No. 96;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHDLVSL 11  
 1:1:11111

Db 11 clgdemdvsl 20

## RESULT 9

AA70374  
 ID AAB70374 standard; Peptide; 26 AA.

AC AAB70374;

DT 02-MAY-2001 (first entry)

DE Bcl BH3 consensus peptide sequence SEQ ID NO:7.

KM Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;  
 KM immunostimulant; neuroprotective; nootropic; antischismatic; vulnerary;  
 KM cytosolic; antiviral; antiarthritic; antiinflammatory; wound healing;

KM immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;  
 KM immunodeficiency disease; neurodegenerative disease; viral infection;  
 KM ischaemic cell death; reperfusion cell death; arthritis; infertility;  
 KM lymphoproliferative condition; inflammation; autoimmune disease.

OS Unidentified.

PN WO200110888-A1.

PD 15-FEB-2001.

PF 30-MAY-2000; 2000WO-US11864.

PR 28-MAY-1999; 99US-0136783.

PA (APOB-) APOPTOSIS TECHNOLOGY INC.

PI Zhou X;

DR WPI: 2001-138734/14.

PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,  
 PT useful for screening for candidate compounds which induce or inhibit  
 PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or  
 PT Ser113.

PS Example 2; Fig 3a; 157pp; English.

CC The present invention describes an isolated or synthetic polypeptide  
 CC (I) comprising a less than full length amino acid sequence of a mutant  
 CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its  
 CC fragment, which contains amino acid substitutions at Ser118 of a human  
 CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine  
 CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,  
 CC nootropic, antischismatic, vulnerary, cytosolic, antiviral,  
 CC antiarthritic, antiinflammatory and immunosuppressive activities, and  
 CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and  
 CC polynucleotides can be used for screening candidate compounds and drugs  
 CC for activity that promote cell survival or apoptosis. Other uses include  
 CC inducing or inhibiting apoptosis in a cell. Candidate compounds  
 CC identified and (mutant) BAD polypeptides are useful in treating  
 CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell  
 CC death, reperfusion cell death, wound healing, cancer, viral infections,  
 CC lymphoproliferative conditions, arthritis, infertility, inflammation and  
 CC autoimmune diseases. The present sequence represents a Bcl-family member  
 CC Bcl3 domain consensus sequence which is used in an example from the  
 CC present invention.

SQ Sequence 26 AA;

Query Match 23 4%; Score 34; DB 22; Length 26;

Best Local Similarity 60.0%; Pred. No. 96;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHDLVSL 11  
 1:1:11111

Db 11 clgdemdvsl 20

## RESULT 10

AA37016  
 ID AAB37016 standard; peptide; 27 AA.

AC AAB37016;

DT 28-FEB-2001 (first entry)

DE Bcl2 polypeptide BH3 domain peptide #16.

KM Cytosolic; neuroprotective; anti-HIV; virucide; cerebroprotective;  
 KM cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;  
 KM apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;

KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;  
 KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;  
 KW stroke; myocardial infarction.

OS Homo sapiens.

PN W0200059526-A1.

PD 12-OCT-2000.

PF 06-APR-2000; 2000MO-US09352.

PR 07-APR-1999; 990US-0128202.

PA (UYDE-) UNIV JEFFERSON THOMAS.

PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;

DR WPI; 2000-679325/66.

PT New peptide conjugates for modulating apoptosis or for inhibiting B  
 cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for  
 treating neurodegenerative disorders, stroke, or cancer -

PS Claim 18; Page 18; 74pp; English.

CC The invention relates to a peptide conjugate having the formula:  
 CC (R-X)-n-peptide where n = 1-10; X = C=O, when the R-X group is attached  
 CC to the N-terminus of the peptide, or a side chain of the peptide where  
 CC the functional group of the side chain is NH<sub>2</sub> or OH; or X = O or NH,  
 CC when the R-X group is attached to the C-terminus of the peptide, or a  
 CC side chain of the peptide, where the side chain functional group is COOH  
 CC or CONH<sub>2</sub>; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl optionally  
 CC monosubstituted with a 1-5C straight or branched chain alkyl group,  
 CC phenyl optionally monosubstituted with a 1-5C straight or branched chain  
 CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples  
 CC of the peptide portion of the conjugate. The peptides represent analogues  
 CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of  
 CC the Bcl-2 domain of the cell death agonist Bad. The peptide conjugate is  
 CC useful for modulating apoptosis in the cells of a subject, or for  
 CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of  
 CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2  
 CC function. In particular, the peptide conjugate is useful for treating a  
 CC subject afflicted with a cancer characterized by cancer cells that  
 CC express Bcl-2. The cancer includes prostate, colorectal, gastric,  
 CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or  
 CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide  
 CC conjugate is also useful for treating disorders characterized by  
 CC increased apoptosis, e.g. neurodegenerative disorders, acquired  
 CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.

XX Sequence 27 AA;

Query Match 23.4%; Score 34; DB 21; Length 27;

Best Local Similarity 60.0%; Pred. No. 1e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHLDVSL 11  
 |||:|||||  
 Db 14 clgdemvsl 23

RESULT 11

AAU015328

ID AAU15328 standard; Peptide; 13 AA.

XX AAU15328;

DT 24-OCT-2001 (first entry)

DE Schizophrenia-associated isoform peptide #213.

XX Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;  
 KW neuroleptic; neuroleptic; diagnostic; neuropsychiatric disorder;  
 KW neurological disorder; neuropathy.

OS Homo sapiens.

PN W0200163293-A2.

PD 30-AUG-2001.

PF 23-FEB-2001; 2001WO-GB00783.

PR 24-FEB-2000; 2000GB-0004415.

PR 28-NOV-2000; 2000US-0750395.

PA (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

PI Herath HMC, Parekh RB, Rohlf C;

DR WPI; 2001-502868/55.

PT Diagnosing and monitoring schizophrenia by detecting the presence of  
 PT Schizophrenia Associated Features and Schizophrenia Associated Protein  
 PT Isoforms in samples of cerebrospinal fluid -

PS Claim 6; Page 33; 160pp; English.

CC The invention relates to methods and compositions for screening,  
 CC diagnosis and prognosis of Schizophrenia. The method involves detecting  
 CC the presence of Schizophrenia (SCH) Associated Features (SFS) and SCH  
 CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,  
 CC immunosay or hybridisation assay, for diagnosing and monitoring SCH,  
 CC studying the effectiveness of treatment and for identifying potential  
 CC therapeutic agents. The method is used for (1) screening or diagnosis of  
 CC SCH and the relative abundance of at least 1 chosen feature correlates  
 CC with the presence or absence of SCH; and (2) monitoring the effect of  
 CC therapy administered to a subject with SCH and the relative abundance of  
 CC at least 1 chosen feature which correlates with the severity of SCH.  
 CC The expression and activity of the SFS, SPIs and related molecules  
 CC (e.g. secondary messengers) are studied to diagnose SCH, monitor the  
 CC progress of the disorder and the effectiveness of treatment and as  
 CC targets to identify and produce potential therapeutic agents for the  
 CC treatment of SCH. The paucity of detectable neuropathic defects  
 CC distinguishes neuropsychiatric disorders such as SCH from neurological  
 CC disorders, where manifestations of anatomical and biochemical changes  
 CC have been identified in many cases. Consequently the identification and  
 CC characterisation of cellular and/or molecular causative defects and  
 CC neuropathies are necessary for improved treatment of neuropsychiatric  
 CC disorders. AAU1514-AAU15762 represent the amino acid sequences of  
 CC schizophrenia-associated isoforms used in the method of the invention.

XX Sequence 13 AA;

Query Match 22.8%; Score 33; DB 22; Length 13;

Best Local Similarity 58.3%; Pred. No. 60;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 SIARSEISVGA 21  
 |||:|:|:|:|  
 Db 1 sldteltdvaee 12

RESULT 12

AAU015338

ID AAU15338 standard; Peptide; 13 AA.

XX AAU15338;

DT 24-OCT-2001 (first entry)

DE Schizophrenia-associated isoform peptide #223.

KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;  
 KW neurological disorder; neuropathy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200163293-A2.  
 PD 30-AUG-2001.  
 XX  
 PD 23-FEB-2001; 2001WO-GB00783.  
 XX  
 PF 24-FEB-2000; 2000GB-0004415.  
 PR 28-NOV-2000; 2000US-0750395.  
 XX  
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PI Herath HMAc, Parekh RB, Rohlf C;  
 DR WPI; 2001-502868/55.  
 XX  
 PT Diagnosing and monitoring Schizophrenia by detecting the presence of  
 PT Schizophrenia Associated Features and Schizophrenia Associated Protein  
 PT Isoforms in samples of cerebrospinal fluid -  
 PS  
 PS Claim 6; Page 33; 160pp; English.  
 XX  
 CC The invention relates to methods and compositions for screening,  
 CC diagnosis and prognosis of Schizophrenia. The method involves detecting  
 CC the presence of Schizophrenia (SCH) Associated Features (SFS) and SCH  
 CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,  
 CC immunosay or hybridisation assay, for diagnosing and monitoring SCH,  
 CC studying the effectiveness of treatments and for identifying potential  
 CC therapeutic agents. The method is used for (1) screening or diagnosis of  
 CC SCH and the relative abundance of at least 1 chosen feature correlates  
 CC with the presence or absence of SCH; and (2) monitoring the effect of  
 CC therapy administered to a subject with SCH and the relative abundance of  
 CC at least 1 chosen feature which correlates with the severity of SCH.  
 CC The expression and activity of the SFS, SPIs and related molecules  
 CC (e.g. secondary messengers) are studied to diagnose SCH, monitor the  
 CC progress of the disorder and the effectiveness of treatment and as  
 CC targets to identify and produce potential therapeutic agents for the  
 CC treatment of SCH. The paucity of detectable neurologic defects  
 CC distinguishes neuropsychiatric disorders such as SCH from neurological  
 CC disorders, where manifestations of anatomical and biochemical changes  
 CC have been identified in many cases. Consequently the identification and  
 CC characterisation of cellular and/or molecular causative defects and  
 CC neuropathies are necessary for improved treatment of neuropsychiatric  
 CC disorders. AA015114-AA015762 represent the amino acid sequences of  
 CC Schizophrenia-associated isoforms used in the method of the invention.  
 CC  
 XX  
 SQ Sequence 13 AA;  
 XX  
 QY  
 DB 10 SLAFSEISWGAE 21  
 DB 1 stldfelfdvaee 12

Query Match 22.8%; Score 33; DB 22; Length 13;  
 Best Local Similarity 58.3%; Pred. No. 60;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 13  
 AAW38432  
 ID AAW38432 standard; Peptide; 15 AA.  
 XX  
 AC AAW38432;  
 XX  
 DT 21-MAY-1998 (first entry)  
 XX  
 DE Human dendritic cell receptor DEC-205 peptide 3 (aa82-96).  
 XX  
 KW Dendritic cell receptor; DEC-205; human; ligand; cell targeting;

KW antigen; toxin; immunosuppressive; therapy; prophylaxis;  
 KW monoclonal antibody.  
 XX  
 OS Synthetic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9745449-A1.  
 PD 04-DEC-1997.  
 XX  
 PD 29-MAY-1997; 97WO-NZ00068.  
 XX  
 PF 29-MAY-1996; 96NZ-0286692.  
 PR 29-MAY-1996; 96NZ-0286692.  
 XX  
 PA (HART/) HART D N J.  
 PI Hart DNJ;  
 DR WPI; 1998-032580/03.  
 XX  
 PT Human dendritic cell receptor, DEC-205 - useful to develop ligands  
 PT for isolation and targeted cell delivery of antigen or toxin  
 XX  
 PS Example 3; Fig 11; 53pp; English.  
 XX  
 CC DEC-205 peptide 3 corresponds to amino acids 82-96 of the human  
 CC dendritic cell receptor DEC-205 (see AAW38429). BALB/c mice have  
 CC been immunised with DEC-205 peptide 1 (see AAW38430) and DEC-205  
 CC peptide 2 (see AAW38431). Spleen cells were subsequently fused with  
 CC NS-1 myeloma cells. A hybridoma, 2F5, was obtained with produced  
 CC monoclonal antibody binding to peptide 1 but not to peptide 2 or  
 CC to control DEC-205 peptide 3. The invention provides isolated  
 CC human DEC-205, its extracellular domain and equivalent fragments,  
 CC and polynucleotides encoding these polypeptides. Further provided  
 CC are ligands (preferably an antibody or antibody binding fragment)  
 CC that bind to human DEC-205, as well as constructs for use in  
 CC prophylaxis and therapy that comprise such a ligand, human DEC-205  
 CC or an extracellular domain coupled to an antigen capable of  
 CC inducing a protective immune response in a patient, or to a toxin  
 CC (e.g. ricin A chain to specifically destroy dendritic cells as part  
 CC of an immunosuppressive process).  
 CC  
 XX  
 SQ Sequence 15 AA;  
 XX  
 QY  
 DB 1 KCLGYHLDVSLAFSEI 16  
 DB 1 kclgy--lditksvnel 14

Query Match 22.8%; Score 33; DB 19; Length 15;  
 Best Local Similarity 43.8%; Pred. No. 72;  
 Matches 7; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

RESULT 14  
 AAB06316  
 ID AAB06316 standard; peptide; 16 AA.  
 XX  
 AC AAB06316;  
 XX  
 DT 03-OCT-2000 (first entry)  
 XX  
 DE Human beta-amyloid precursor protein beta-secretase site.  
 XX  
 DE Human; beta-amyloid precursor protein; beta-APP; beta-secretase;  
 KW subtilisin-kexin isoenzyme 1; SKI-1;  
 KW pro-brain-derived neurotrophic factor; proBDNF; antileptemic;  
 KW cytosolic; vasotropic; SKI-1 inhibitor; hypercholesterolaemia;  
 KW liver steatosis; Ras-dependent cancer; restenosis;  
 KW amyloid protein formation.  
 XX  
 XX Mammalia.  
 OS  
 XX

FH Key Location/Qualifiers  
 FT Cleavage-site 8...9  
 PN WO200026348-A2.  
 PD 11-MAY-2000.  
 XX  
 PF 04-NOV-1999; 99WO-CA01058.  
 XX  
 PR 04-NOV-1998; 98CA-2249648.  
 XX  
 PA (RECL-) INST RECH CLINIQUES MONTREAL.  
 PI Seidah N, Chretien M, Marcinkiewicz M, Laaksonen R, Davignon J;  
 DR WPI: 2000-365601/31.  
 XX  
 PT Novel soluble proteic fragment of subtilisin-kexin isoenzyme for  
 PT producing a polypeptide useful for treating hypercholesterolemia, liver  
 PT steatosis and amyloidosis, comprises a specific amino acid sequence -  
 XX  
 PS Example 4; Page 51; 119pp; English.  
 CC  
 CC The present sequence is the beta-secretase site of human beta-amyloid  
 CC precursor protein (beta-APP). The sequence may be cleaved by a mammalian  
 CC membrane-bound proteinase. Peptides which bind to and are cleaved by  
 CC SKI-1 may be used for monitoring SKI-1 activity, for screening inhibitors  
 CC of SKI-1 activity, or for screening enhancers of SKI-1 activity. Proteic  
 CC fragments of SKI-1 which bind to the SKI-1 catalytic site may be used as  
 CC inhibitors of SKI-1 activity. They may be used to treat diseases  
 CC involving overexpression of SKI-1 or SKI-1 substrate. Such diseases  
 CC include hypercholesterolemia, high levels of fatty acids, lipids or  
 CC fatty acid pyrophosphate, liver steatosis, Ras-dependent cancer, restenosis  
 CC and amyloid protein formation.  
 XX  
 SQ Sequence 16 AA;  
 XX  
 Query Match 22.8%; Score 33; DB 21; Length 16;  
 Best Local Similarity 50.0%; Pred. No. 78;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 14 SEISVGAENKND 25  
 DB 4 sevnidaefrhd 15  
 XX  
 RESULT 15  
 AAE00609  
 ID AAE00609 standard; peptide; 18 AA.  
 XX  
 AC AAE00609;  
 XX  
 DT 02-JUL-2001 (first entry)  
 XX  
 DE Beta-amyloid precursor protein beta-secretase cleavage site mutant.  
 XX  
 KW Beta-amyloid precursor protein; beta-APP; caspase; beta-secretase;  
 KW cysteine protease; apoptosis; caspase expression cassette; metastasis;  
 KW tumour; cathepsin B; urokinase; proliferation; gene therapy;  
 KW interdomain linker; cleavage site; Alzheimer's disease; mutant; mutein.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200129232-A2.  
 PD 26-APR-2001.  
 XX  
 PF 19-OCT-2000; 2000WO-US28941.  
 XX  
 PR 20-OCT-1999; 99US-0160559.  
 PR 14-AUG-2000; 2000US-0225564.

XX  
 PA (SCIO-) SCIOS INC.  
 XX  
 PI Cordell B, Li Y;  
 XX  
 DR WPI: 2001-290920/30.  
 XX  
 PT Novel fusion polypeptide comprising first and second caspase subunit  
 PT separated by cleavage site not associated in nature with caspase  
 PT subunit, useful for cloning gene encoding enzymes involved in  
 PT proteolytic cleavage -  
 XX  
 PS Example 2; Page 26; 116pp; English.  
 XX  
 CC The present sequence is a beta-secretase cleavage site of beta-amyloid  
 CC precursor protein (beta-APP) mutant found in certain families of  
 CC autosomal dominant form of Alzheimer's disease. This sequence is used to  
 CC construct a chimeric cassette comprising human caspase-3 with interdomain  
 CC linker replaced by this sequence. This modified caspase-3 plays a pivotal  
 CC role in Alzheimer's disease. Caspases are a family of cysteine proteases,  
 CC that participate in the initiation and execution of apoptosis.  
 CC The present invention relates to a method for functional cloning of genes  
 CC encoding proteins or enzymes involved in proteolytic cleavage. The  
 CC invention is based on the use of caspase expression cassettes comprising  
 CC the coding sequence of a proteolytic cleavage site flanked by sequences  
 CC and a second caspase subunits. A fusion polypeptide comprising a first  
 CC in nature, is useful for cloning gene encoding enzymes involved in  
 CC proteolytic cleavage. An expression cassette containing fusion  
 CC polypeptide is used to identify a mutant cell line deficient in an  
 CC enzyme of interest and is also useful for diagnosis and suppression of  
 CC overexpression or metastases of a tumour cell characterised by  
 CC selective expression of a polypeptide (e.g. Cathepsin B or urokinase,  
 CC polypeptide is expressed in the tumour cells). DNA encoding fusion  
 XX  
 SQ Sequence 18 AA;  
 XX  
 Query Match 22.8%; Score 33; DB 22; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 90;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 14 SEISVGAENKND 25  
 DB 6 sevnidaefrhd 17  
 XX  
 RESULT 16  
 AAE00611  
 ID AAE00611 standard; peptide; 19 AA.  
 XX  
 AC AAE00611;  
 XX  
 DT 02-JUL-2001 (first entry)  
 XX  
 DE Amyloid precursor protein (APP) beta-secretase cleavage site.  
 XX  
 KW Amyloid precursor protein; APP; caspase; beta-secretase; cleavage site;  
 KW cysteine protease; apoptosis; caspase expression cassette; metastasis;  
 KW tumour; cathepsin B; urokinase; proliferation; gene therapy;  
 KW interdomain linker; Alzheimer's disease.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200129232-A2.  
 PD 26-APR-2001.  
 XX  
 PF 19-OCT-2000; 2000WO-US28941.  
 XX  
 PR 20-OCT-1999; 99US-0160559.  
 PR 14-AUG-2000; 2000US-0225564.

```

XX (SCIO-) SCIOS INC.
PA
XX
PI Cordell B, Li Y;
XX
DR WPI; 2001-290920/30.
XX
PT Novel fusion polypeptide comprising first and second caspase subunit
PT separated by cleavage site not associated in nature with caspase
PT subunit, useful for cloning gene encoding enzymes involved in
PT proteolytic cleavage
XX
PS Disclosure; Fig 28A; 116pp; English.
XX
XX The present amino acid sequence is a beta-secretase cleavage site of
CC an amyloid precursor protein (APP). This sequence is used to construct
CC an artificially engineered chimeric cassette comprising human caspase-3
CC with interdomain linker replaced by Swedish mutant beta-secretase
CC cleavage site. This modified caspase-3 plays a pivotal role in
CC Alzheimer's disease. Caspases are a family of cysteine proteases, that
CC participate in the initiation and execution of apoptosis.
CC The present invention relates to a method for functional cloning of genes
CC encoding proteins or enzymes involved in proteolytic cleavage. The
CC invention is based on the use of caspase expression cassettes comprising
CC the coding sequence of a proteolytic cleavage site flanked by sequences
CC encoding two caspase subunits. A fusion polypeptide comprising a first
CC and a second caspase subunit, separated by a cleavage site not associated
CC in nature, is useful for cloning gene encoding enzymes involved in
CC proteolytic cleavage. An expression cassette containing fusion
CC polypeptide is used to identify a mutant cell line deficient in an
CC enzyme of interest and is also useful for diagnosis and suppression of
CC proliferation or metastases of a tumour cell characterised by
CC overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC selectively expressed in the tumour cells). DNA encoding fusion
CC polypeptide is used in gene therapy.
XX
SQ Sequence 19 AA;

Query Match 22.8%; Score 33; DB 22; Length 19;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25
|:::| | |
Db 7 sevnldaeefrhd 18

RESULT 17
AAy69714
ID AAY69714 standard; peptide; 20 AA.
XX
AC AAY69714;
XX
DT 11-APR-2000 (first entry)
XX
DE beta-APP alpha-secretase substrate [NL]-APP(-10,+10).
XX
KM Neurotrophic; neuroprotective; beta-amyloid precursor protein; metabolism;
KM cleavage site; beta-secretase; neurodegenerative disease;
KM Alzheimer's disease.
XX
XX Homo sapiens.
XX OS Synthetic.
XX PN W09964587-A1.
XX PD 16-DEC-1999.
XX PF 04-JUN-1999; 99MO-FR01326.
XX PR 05-JUN-1998; 98FR-0007068.
XX PR 31-MAR-1999; 99US-0122599.

```

```

XX (RHON ) RHONE-POULENC RORER SA.
PA (UYPA-) UNIV CURIE PARIS VI P & M.
XX
XX Rholam M, Munoz-Gimenez N, Moutaouakil M, Cohen P, Bertrand P;
XX
DR WPI; 2000-097537/08.
XX
PT Polypeptide with beta-secretase activity, specific for wild-type
PT amyloid precursor protein, useful in treating Alzheimer's disease
XX
PS Example 3; Page 24; 44pp; French.
XX
XX Peptides AAY69702-Y69718 represent synthetic peptide substrates for a
CC novel polypeptide with beta-secretase activity that can cleave
CC specifically the natural beta-amyloid precursor protein (bAPP). Normal
CC cleavage of the protein occurs between amino acids Met596-Asp597 and
CC Val636-Ile637 (positions 4-5 and 44-45 of AAY69701). The novel
CC polypeptide is used to identify agents that interact specifically with
CC it. These agents regulate metabolism of APP, particularly they slow down
CC or reduce production of beta-amyloid, so can be used to treat
CC neurodegenerative diseases, particularly Alzheimer's disease.
XX
SQ Sequence 20 AA;

Query Match 22.8%; Score 33; DB 21; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25
|:::| | |
Db 6 sevnldaeefrhd 17

RESULT 18
AAy69716
ID AAY69716 standard; peptide; 20 AA.
XX
AC AAY69716;
XX
DT 11-APR-2000 (first entry)
XX
DE beta-APP alpha-secretase substrate [NM]-APP(-10,+10).
XX
KM Neurotrophic; neuroprotective; beta-amyloid precursor protein; metabolism;
KM cleavage site; beta-secretase; neurodegenerative disease;
KM Alzheimer's disease.
XX
XX Homo sapiens.
XX OS Synthetic.
XX PN W09964587-A1.
XX PD 16-DEC-1999.
XX PF 04-JUN-1999; 99MO-FR01326.
XX PR 05-JUN-1998; 98FR-0007068.
XX PR 31-MAR-1999; 99US-0122599.
XX
XX (RHON ) RHONE-POULENC RORER SA.
XX PA (UYPA-) UNIV CURIE PARIS VI P & M.
XX PI Rholam M, Munoz-Gimenez N, Moutaouakil M, Cohen P, Bertrand P;
XX DR WPI; 2000-097537/08.
XX
PT Polypeptide with beta-secretase activity, specific for wild-type
PT amyloid precursor protein, useful in treating Alzheimer's disease
XX
PS Example 3; Page 24; 44pp; French.

```

CC Peptides AAY69702-Y69718 represent synthetic peptide substrates for a  
 CC novel polypeptide with beta-secretase activity that can cleave  
 CC specifically the natural beta-amyloid precursor protein (bAPP). Normal  
 CC cleavage of the protein occurs between amino acids Met596-Asp597 and  
 CC Val636-Ile637 (positions 4-5 and 44-45 of AAY69701). The novel  
 CC polypeptide is used to identify agents that interact specifically with  
 CC it. These agents regulate metabolism of APP, particularly they slow down  
 CC or reduce production of beta-amyloid, so can be used to treat  
 CC neurodegenerative diseases, particularly Alzheimer's disease.  
 XX Sequence 20 AA;

Query Match 22.8%; Score 33; DB 21; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
 ||::||| |  
 Db 6 sevnmdaefrhd 17

## RESULT 19

AAW08361  
 ID AAW08361 standard; peptide; 21 AA.

AC AAW08361;

DT 05-SEP-1997 (first entry)

DE Beta-secretase substrate #3.

KW Beta-cleavage site; beta amyloid precursor protein; APP; beta-secretase;  
 KW alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.

OS Synthetic.

PN W09640885-A2.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US09985.

PR 07-JUN-1995; 95US-0485152.

PR 07-JUN-1995; 95US-0480498.

PA (ATHE-) ATHENA NEUROSCIENCES INC.

PI Anderson JP, Chrysler SMS, Jacobson-croak KL, Kelm PS;  
 PI Mcconlogue LC, Sinha S, Tan H;

DR WPI; 1997-052304/05.

PT Beta-secretase which specifically cleaves beta-amyloid precursor  
 PT protein - useful to screen for inhibitors useful in treatment of  
 XX Alzheimer's disease

PS Disclosure; Page 45; 92pp; English.

AAW08359-W08362 represent substrates for the enzyme of the invention.  
 CC The enzyme of the invention is beta-secretase, and specifically cleaves  
 CC beta-amyloid precursor protein (beta-APP). Normal processing of beta-APP  
 CC is thought to occur via cleavage between residues 16 and 17 of the  
 CC beta-amyloid peptide region by an alpha-secretase. Pathogenic processing  
 CC is thought to occur by beta-secretase cleavage of beta-APP.  
 CC Beta-secretase activity can be detected and measured using a method of  
 CC the invention, which detects at least one of the beta-secretase cleavage  
 CC products formed on cleavage. The method can be used to determine whether  
 CC a test substance inhibits proteolytic cleavage, by beta-secretase, of  
 CC beta-APP. Compounds effective to at least partially inhibit  
 CC beta-secretase activity can be used to inhibit cleavage of beta-APP in  
 CC cells or mammalian hosts. Isolation and purification of beta-secretase  
 CC will permit chemical modelling of a critical event in the pathology of

CC Alzheimer's disease.  
 XX Sequence 21 AA;  
 SQ

Query Match 22.8%; Score 33; DB 18; Length 21;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
 ||::||| |  
 Db 1 sevnldaeefrhd 12

## RESULT 20

AAW82188  
 ID AAW82188 standard; peptide; 21 AA.

AC AAW82188;

DT 18-FEB-1999 (first entry)

DE Fluorogenic protease indicator native amyloid peptide.

KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;  
 KW conformation change.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 3

FT /label= Alb

FT /note= "alpha-aminobutyric acid, labelled as

FT Modified-site 4 amino acid B in the specification"

FT /note= "epsilon-aminocaproic acid, labelled as

FT Modified-site 16 amino acid J in the specification"

FT /note= "epsilon-aminocaproic acid, labelled as

PN W09837226-A1.

PD 27-AUG-1998.

PF 20-FEB-1998; 98WO-US03000.

PR 20-FEB-1997; 97US-0802981.

PA (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

DR WPI; 1998-467579/40.

PT New fluorogenic compositions - containing 2 fluorophores separated  
 PT by a peptide comprising a protease binding site, used for detecting  
 XX protease activity in samples.

PS Disclosure; Page 29; 90pp; English.

AAW82023-W82240 are peptides used in the construction of a fluorogenic  
 CC composition which is used for the detection of protease activity in  
 CC biological samples. The products can be used for the detection of  
 CC conformational changes in nucleic acids, oligosaccharides,  
 CC polysaccharides, proteins, peptides, lipids, phospholipids, glycolipids,  
 CC glycoproteins, steroids or polymers. In addition, attachment of a  
 CC hydrophobic group to a molecule can be used to enhance uptake of a  
 CC The composition is composed of P = peptide comprising a protease binding  
 CC site for the protease, F1, F2 peptides = fluorophores where F1 is  
 CC attached to the amino terminal amino acid and F2 is attached to the  
 CC carboxyl terminal amino acid and S1, S2 peptides = when present, are  
 CC peptide spacers where S1, when present, is attached to the amino terminal



CC acid, and S2, when present, is attached to the carboxyl terminal amino  
 CC acid.  
 XX  
 SQ Sequence 21 AA;

Query Match 22.8%; Score 33; DB 19; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 SEISVGAERN--KDD 26  
 ||: : ||| |||  
 Db 6 sevkmdaefgpxkdd 20

RESULT 21  
 AAY33755  
 ID AAY33755 standard; Protein; 21 AA.  
 XX  
 AC AAY33755;  
 XX  
 DT 09-NOV-1999 (first entry)  
 XX  
 DE Synthetic oligopeptide 5-16'SW.  
 XX  
 KM Beta-secretase; beta-amyloid protein precursor; APP; Down's syndrome;  
 KM Alzheimer's disease; measure activity; cleavage site.  
 XX  
 OS Synthetic.  
 XX  
 PN US5942400-A.  
 XX  
 PD 24-AUG-1999.  
 XX  
 PE 07-JUN-1996; 96US-0659984.  
 XX  
 PR 07-JUN-1996; 96US-0659984.  
 PR 07-JUN-1995; 95US-0480498.  
 PR 07-JUN-1995; 95US-0485152.  
 XX  
 PA (ELAN-) ELAN PHARM INC.  
 PI Anderson JP, Jacobson-Croak KL, Sinha S;  
 DR WPI; 1999-517417/43.  
 XX  
 PT A method for detecting human beta-secretase cleavage of polypeptides  
 PT useful for identifying beta-secretase inhibitors  
 XX  
 PS Examples; Column 30; 43pp; English.  
 XX  
 CC Sequences AAY33752-Y33756 are synthetic oligopeptides used for measuring  
 CC the activity of beta-secretase (AAY33741). Beta-secretase is capable of  
 CC cleaving beta-amyloid protein precursor (APP) (AAY33742). These  
 CC synthetic peptides contain the cleavage site of APP. Beta-secretase and  
 CC APP are used in a method for detecting human beta-secretase cleavage of  
 CC polypeptides and for identifying beta-secretase inhibitors. Inhibition  
 CC of beta-secretase activity would be useful for chemical modelling of a  
 CC critical event in the pathology of Alzheimer's disease. Inhibitors of  
 CC beta-secretase would be useful for the prevention and treatment of  
 CC Alzheimer's disease and Down's Syndrome.  
 CC  
 SQ Sequence 21 AA;

Query Match 22.8%; Score 33; DB 20; Length 21;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 14 SEISVGAERNKD 25  
 ||::: ||| |  
 Db 1 sevnldaefrhd 12

RESULT 22  
 AAG73205  
 ID AAG73205 standard; Peptide; 21 AA.  
 XX

AC AAG73205;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Protease binding site #139.  
 XX  
 KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KM viral infection; cancer metastasis; emphysema; arthritis;  
 KM thrombosis; haemophilia.  
 XX  
 OS Synthetic.

XX  
 FH Key Location/Qualifiers  
 FT Modified-site 3 /label= Alb  
 FT FT /note= "2-aminoisobutyric acid"  
 XX

PN MO200118238-A1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-US24882.  
 XX  
 PR 10-SEP-1999; 99US-0394019.  
 XX

PA (ONCO-) ONCOIMMUNIN INC.  
 XX  
 PI Komoriya A, Packard BS;  
 XX  
 DR WPI; 2001-389573/41.  
 XX

XX  
 PT New fluorogenic compositions whose fluorescence level increases in the  
 PT presence of active proteases, useful for detecting and localizing  
 PT protease activity in biological samples, particularly in frozen tissue  
 PT samples  
 XX  
 PS Disclosure; Page 27; 86pp; English.

XX  
 CC The present invention describes fluorogenic compositions which can be  
 CC used for the detection of protease activity. This can be useful as an  
 CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,  
 CC thrombosis and arthritis. The fluorogenic compositions comprise a  
 CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The  
 CC peptide is cleaved by a protease and the fluorophores can then be  
 CC detected. The present sequence is one of the peptides described in the  
 CC exemplification of the invention.  
 XX

SQ Sequence 21 AA;

Query Match 22.8%; Score 33; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 SEISVGAERN--KDD 26  
 ||: : ||| |||  
 Db 6 sevkldaefgpxkdd 20

RESULT 23  
 AAG73206  
 ID AAG73206 standard; Peptide; 21 AA.  
 XX  
 AC AAG73206;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Protease binding site #140.

```

XX  Protease detection; peptide cleavage; enzyme activity; fluorogenic;
KW viral infection; cancer metastasis; emphysema; arthritis;
XX thrombosis; haemophilia.
OS Synthetic.
XX Key Location/Qualifiers
FT Modified-site 3 /label= Aib
FT /note= "2-aminoisobutyric acid"
PN WC200118238-A1.
XX 15-MAR-2001.
PD 11-SEP-2000; 2000WO-US24882.
XX 10-SEP-1999; 99US-0394019.
PR (ONCO-) ONCOIMMUNIN INC.
XX (ONCO-) ONCOIMMUNIN INC.
PA Komoriya A, Packard BS;
XX WPI; 2001-389573/41.
DR 2001-389573/41.
XX
PI New fluorogenic compositions whose fluorescence level increases in the
PT presence of active proteases, useful for detecting and localizing
PT protease activity in biological samples, particularly in frozen tissue
XX samples -
XX
PS Disclosure; Page 27; 86pp; English.
XX
CC The present invention describes fluorogenic compositions which can be
CC used for the detection of protease activity. This can be useful as an
CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
CC thrombosis and arthritis. The fluorogenic compositions comprise a
CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
CC peptide is cleaved by a protease and the fluorophores can then be
CC detected. The present sequence is one of the peptides described in the
CC exemplification of the invention.
XX
SQ Sequence 21 AA;

```

Query Match 22.8%; Score 33; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. NO. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 14 SEISVGAEFN--KDD 26  
 ||: ||| |||  
 Db 6 sevkmdaefgcpkd 20

RESULT 24  
 AAG73316  
 ID AAG73316 standard; Peptide; 21 AA.  
 AC AAG73316;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Protease indicator compound peptide #45.  
 XX  
 KW Protease detection; peptide cleavage; enzyme activity; fluorogenic;  
 KW viral infection; cancer metastasis; emphysema; arthritis;  
 XX thrombosis; haemophilia.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER  
 FT

```

FT Modified-site /note= "modified by fm"
FT 4 /label= OTHER
FT /note= "designated J in the specification"
FT Misc-difference 7 /note= "D-form residue"
FT Misc-difference 10 /note= "D-form residue"
FT Misc-difference 14 /note= "D-form residue"
FT Modified-site 16 /note= "D-form residue"
FT /label= OTHER
FT /note= "designated C5 in the specification"
PN WC200118238-A1.
XX 15-MAR-2001.
PD 11-SEP-2000; 2000WO-US24882.
XX 10-SEP-1999; 99US-0394019.
PR (ONCO-) ONCOIMMUNIN INC.
XX (ONCO-) ONCOIMMUNIN INC.
PA Komoriya A, Packard BS;
XX WPI; 2001-389573/41.
DR 2001-389573/41.
XX
PI New fluorogenic compositions whose fluorescence level increases in the
PT presence of active proteases, useful for detecting and localizing
PT protease activity in biological samples, particularly in frozen tissue
XX samples -
XX
PS Claim 4; Page 71; 86pp; English.
XX
CC The present invention describes fluorogenic compositions which can be
CC used for the detection of protease activity. This can be useful as an
CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
CC thrombosis and arthritis. The fluorogenic compositions comprise a
CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
CC peptide is cleaved by a protease and the fluorophores can then be
CC detected. The present sequence is one of the peptides described in the
CC exemplification of the invention.
XX
SQ Sequence 21 AA;

```

Query Match 22.8%; Score 33; DB 22; Length 21;  
 Best Local Similarity 53.3%; Pred. NO. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 14 SEISVGAEFN--KDD 26  
 ||: ||| |||  
 Db 6 sevkmdaefgcpkd 20

RESULT 25  
 AAB47265  
 ID AAB47265 standard; Peptide; 21 AA.  
 AC AAB47265;  
 XX  
 DT 18-JUL-2001 (first entry)  
 XX  
 DE Peptide 5-16'SW, for used in beta-secretase assay.  
 XX  
 KW Beta-secretase; isotype; beta-amyloid precursor protein; app;  
 KW beta-amyloid peptide; beta-AP; Alzheimer's disease; Downs syndrome;  
 XX hCHMA-D; Swedish mutation; maltose binding protein; MBP.  
 OS Synthetic.  
 XX US6221645-B1.  
 PN

```

XX 24-APR-2001.
PD
XX
XX 07-JUN-1996; 96US-0660531.
PF
XX 07-JUN-1995; 95US-0480498.
PR
XX
XX (ELAN-) ELAN PHARM INC.
PA
XX
XX Chrysler SMS, Sinha S, Keim PS, Anderson JP, Tan H, McConlogue LC;
XX
XX WPI; 2001-315578/33.
DR
XX
XX Novel antibody that specifically binds native beta-secretase protein,
PR useful for raising anti-idiotypic antibodies and for detecting or
PT diagnosing pathological conditions related to presence of respective
PT antigens
XX
XX Example: Column 30; 42pp; English.
PS
XX
XX The sequences given in ABA47262-67 represent synthetic peptides
CC containing the cleavage sites derived from wild-type beta-amyloid
CC precursor protein (APP). These peptides were used in assays utilizing
CC partially purified beta-secretase to identify beta-secretase inhibitors.
CC Beta-secretase is thought to be responsible for the pathogenic
CC processing of APP to form beta amyloid peptide (beta-Ap) in beta-Ap
CC related conditions, e.g. Alzheimer's disease, Downs syndrome, HCHWA-D
CC etc. Beta-secretase has a molecular weight of 260-300 KD and will bind
CC to wheat germ agglutinin but not to concanavalin A. Beta-secretase
CC will cleave both the wild type and the Swedish mutation of APP.
XX
XX Sequence 21 AA;
SQ

```

```

Query Match 22.8%; Score 33; DB 22; Length 21;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 14 SEISVGAEFNKD 25
| | : : | | | |
Db 1 sevnldaefrhd 12

```

Search completed: January 6, 2002, 09:12:21  
 Job time: 434 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2002, 09:06:02 ; Search time 16.63 Seconds  
(without alignments)  
36.536 Million cell updates/sec

Title: US-09-020-393b-18

Perfect score: 145  
Sequence: 1 KCLGYHLDVSLAFSEISVGAEFRKDDC 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 127597

Minimum DB seq length: 0  
Maximum DB seq length: 27

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfillsl.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                   | Description        |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1          | 38    | 26.2        | 16     | 5 PCT-US94-07043A-1  | Sequence 1, Appl   |
| 2          | 34    | 23.4        | 20     | 1 US-08-614-935-17   | Sequence 17, Appl  |
| 3          | 34    | 23.4        | 20     | 3 US-09-130-287-17   | Sequence 17, Appl  |
| 4          | 34    | 23.4        | 20     | 4 US-09-236-385A-41  | Sequence 41, Appl  |
| 5          | 34    | 23.4        | 21     | 3 US-08-802-981-112  | Sequence 112, Appl |
| 6          | 33    | 22.8        | 21     | 3 US-08-659-984A-18  | Sequence 18, Appl  |
| 7          | 33    | 22.8        | 21     | 3 US-08-802-981-113  | Sequence 113, Appl |
| 8          | 33    | 22.8        | 21     | 3 US-08-802-981-114  | Sequence 114, Appl |
| 9          | 33    | 22.8        | 21     | 4 US-08-660-531-18   | Sequence 18, Appl  |
| 10         | 32    | 22.1        | 22     | 2 US-08-313-200-8    | Sequence 8, Appl   |
| 11         | 32    | 22.1        | 27     | 1 US-08-141-324-11   | Sequence 11, Appl  |
| 12         | 32    | 22.1        | 27     | 1 US-08-141-324-12   | Sequence 12, Appl  |
| 13         | 32    | 22.1        | 27     | 1 US-08-541-902-11   | Sequence 11, Appl  |
| 14         | 32    | 22.1        | 27     | 1 US-08-541-902-12   | Sequence 12, Appl  |
| 15         | 31    | 21.4        | 19     | 1 US-08-324-301-3    | Sequence 3, Appl   |
| 16         | 30    | 20.7        | 13     | 2 US-08-706-741B-44  | Sequence 44, Appl  |
| 17         | 30    | 20.7        | 13     | 2 US-08-924-695A-44  | Sequence 44, Appl  |
| 18         | 30    | 20.7        | 15     | 1 US-08-440-391-7    | Sequence 7, Appl   |
| 19         | 30    | 20.7        | 15     | 1 US-08-440-391-32   | Sequence 32, Appl  |
| 20         | 30    | 20.7        | 15     | 2 US-08-248-839C-152 | Sequence 152, Appl |
| 21         | 30    | 20.7        | 15     | 2 US-08-908-597A-7   | Sequence 7, Appl   |
| 22         | 30    | 20.7        | 15     | 2 US-08-908-597A-32  | Sequence 32, Appl  |
| 23         | 30    | 20.7        | 15     | 4 US-09-236-385A-7   | Sequence 7, Appl   |
| 24         | 30    | 20.7        | 15     | 4 US-09-236-385A-32  | Sequence 32, Appl  |
| 25         | 30    | 20.7        | 15     | 5 PCT-US96-06122-7   | Sequence 7, Appl   |
| 26         | 30    | 20.7        | 15     | 5 PCT-US96-06122-32  | Sequence 32, Appl  |
| 27         | 30    | 20.7        | 18     | 5 PCT-US94-01234-21  | Sequence 21, Appl  |

|    |    |      |    |                     |                   |
|----|----|------|----|---------------------|-------------------|
| 28 | 29 | 20.0 | 16 | 4 US-09-171-705-41  | Sequence 41, Appl |
| 29 | 29 | 20.0 | 24 | 1 US-08-461-597-4   | Sequence 4, Appl  |
| 30 | 29 | 20.0 | 24 | 2 US-08-535-298-4   | Sequence 4, Appl  |
| 31 | 29 | 20.0 | 24 | 5 PCT-US94-05569A-4 | Sequence 4, Appl  |
| 32 | 29 | 20.0 | 24 | 5 PCT-US94-05569A-4 | Sequence 4, Appl  |
| 33 | 29 | 20.0 | 26 | 1 US-08-475-989-25  | Sequence 25, Appl |
| 34 | 29 | 20.0 | 26 | 2 US-08-475-985-25  | Sequence 25, Appl |
| 35 | 29 | 20.0 | 26 | 3 US-08-256-839-25  | Sequence 25, Appl |
| 36 | 28 | 19.3 | 9  | 3 US-08-802-981-219 | Sequence 19, Appl |
| 37 | 28 | 19.3 | 10 | 2 US-08-659-984A-19 | Sequence 19, Appl |
| 38 | 28 | 19.3 | 10 | 4 US-08-660-531-19  | Sequence 3, Appl  |
| 39 | 28 | 19.3 | 11 | 5 PCT-US94-07043A-3 | Sequence 3, Appl  |
| 40 | 28 | 19.3 | 19 | 1 US-07-678-974D-67 | Sequence 67, Appl |
| 41 | 28 | 19.3 | 19 | 2 US-08-985-090-8   | Sequence 8, Appl  |
| 42 | 28 | 19.3 | 19 | 3 US-09-165-543-8   | Sequence 8, Appl  |
| 43 | 28 | 19.3 | 19 | 4 US-09-165-543-15  | Sequence 15, Appl |
| 44 | 28 | 19.3 | 20 | 4 US-09-007-905-59  | Sequence 59, Appl |
| 45 | 28 | 19.3 | 20 | 4 US-09-007-905-64  | Sequence 64, Appl |

#### ALIGNMENTS

RESULT 1  
PCT-US94-07043A-1  
Sequence 1, Application PC/TUS9407043A  
GENERAL INFORMATION:  
APPLICANT: Tamburini, Paul P.; Benz, Ginter; H. Bich,  
APPLICANT: Dieter; Dreyer, Robert N.; Koenig, Gerhard  
TITLE OF INVENTION: CATHEPSIN D IS AN AMYLOIDGENIC  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Miles Inc.  
STREET: 400 Morgan Lane  
CITY: West Haven  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06516  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage  
COMPUTER: Sharp PC 4600  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07043A  
FILING DATE: June 21, 1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10889  
FILING DATE: November 12, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/995,660  
FILING DATE: December 16, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/880,914  
FILING DATE: May 11, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pamela A. Simonon  
REGISTRATION NUMBER: 31,060  
REFERENCE/DOCKET NUMBER: MTI 224.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 937-2795  
TELEFAX: (203) 937-2795  
INFORMATION FOR SEQ. ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US94-07043A-1

Query Match 26.2% Score 38; DB 5; Length 16;

Best Local Similarity 53.8%; Pred. No. 4 1;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 14 SEISGAEFNKDD 26  
Db 2 SEVKMDAEFRHDD 14

RESULT 2  
US-08-614-935-17  
Sequence 17, Application US/08614935  
Patent No. 5804201  
GENERAL INFORMATION:  
APPLICANT: King, Te P.  
TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES OF VESPID  
NUMBER OF SEQUENCES: 81  
CURRENT APPLICATION DATA:  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-156  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-614-935-17

Query Match 23.4%; Score 34; DB 1; Length 20;  
Best Local Similarity 71.4%; Pred. No. 25;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 21 EFNKDDC 27  
Db 5 DFNHDDC 11

RESULT 3  
US-09-130-287-17  
Sequence 17, Application US/09130287  
Patent No. 6106844  
GENERAL INFORMATION:  
APPLICANT: King, Te P.  
TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES OF VESPID  
NUMBER OF SEQUENCES: 81  
CURRENT APPLICATION DATA:  
FILING DATE: 25-Jan-1999

ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/614,935  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-156 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-09-130-287-17

Query Match 23.4%; Score 34; DB 3; Length 20;  
Best Local Similarity 71.4%; Pred. No. 25;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 21 EFNKDDC 27  
Db 5 DFNHDDC 11

RESULT 4  
US-09-236-385A-41  
Sequence 41, Application US/09236385A  
Patent No. 6221615  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
LUTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 41  
CURRENT APPLICATION DATA:  
FILING DATE: 25-Jan-1999

```

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 41
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 41
US-09-236-385A-41

Query Match      23.4%; Score 34; DB 4; Length 20;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 CLGYHLDVSL 11
        1:1 :||||
Db       11 CIGDEMDVSL 20

RESULT      5
US-08-802-981-112
Sequence 112, Application US/08802981
Patent No. 6037137
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of Enzyme
NUMBER OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-000300US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product= "A1b"
FEATURE:
NAME/KEY: Modified-site

```

```

: LOCATION: 4 /product= "Acp"
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 16
: OTHER INFORMATION: /product= "Acp"
US-08-802-981-112

Query Match 23.4%; Score 34; DB 3; Length 21;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 14 SEISYGAEFN-KDD 26
|::: ||| |||
Db 6 SEVNLDAEFGXPKDD 20

RESULT 6
US-08-659-984A-18
: Sequence 18, Application US/08659984A
: Patent No. 5942400
: GENERAL INFORMATION:
: APPLICANT: Anderson, John P.
: APPLICANT: Sinha, Sukanto
: APPLICANT: Jacobson-Croak, Kirsten L.
: TITLE OF INVENTION: Assays for Detecting Beta-Secretase
: TITLE OF INVENTION: Inhibition
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Ctr., 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/659,984A
: FILING DATE: 07-JUN-1996
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/485,152
: FILING DATE: 07-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Heslin, James M.
: REGISTRATION NUMBER: 29,541
: REFERENCE/DOCKET NUMBER: 15270-0028100S
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-326-2400
: TELEFAX: 415-326-2422
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-659-984A-18

Query Match 22.8%; Score 33; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISYGAEFNKD 25
|::: ||| |
Db 1 SEVNLDAEFRHD 12

```

RESULT 7  
US-08-802-981-113  
; Sequence 113, Application US/08802981  
; Patent No. 6037137  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; APPLICANT: Packard, Beverly S.  
; TITLE OF INVENTION: Compositions for the Detection of Enzyme  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/802,981  
; FILING DATE: 20-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunter, Tom  
; REGISTRATION NUMBER: 38,498  
; REFERENCE/DOCKET NUMBER: 016865-000300US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 113:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; MOLECULE TYPE: linear  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /product= "Aib"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /product= "Acp"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 16  
; OTHER INFORMATION: /product= "Acp"  
US-08-802-981-113

Query Match 22.8%; Score 33; DB 3; Length 21;  
Best Local Similarity 53.3%; Pred. No. 39;  
Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

OY 14 SEISYGAEFN--KDD 26  
||: ||| |||  
Db 6 SEVKDAEFGXPKDD 20

RESULT 8  
US-08-802-981-114  
; Sequence 114, Application US/08802981  
; Patent No. 6037137  
; GENERAL INFORMATION:  
; APPLICANT: Komoriya, Akira  
; APPLICANT: Packard, Beverly S.  
; TITLE OF INVENTION: Compositions for the Detection of Enzyme

; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/802,981  
; FILING DATE: 20-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunter, Tom  
; REGISTRATION NUMBER: 38,498  
; REFERENCE/DOCKET NUMBER: 016865-000300US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; MOLECULE TYPE: linear  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /product= "Aib"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /product= "Acp"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 16  
; OTHER INFORMATION: /product= "Acp"  
US-08-802-981-114

Query Match 22.8%; Score 33; DB 3; Length 21;  
Best Local Similarity 53.3%; Pred. No. 39;  
Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

OY 14 SEISYGAEFN--KDD 26  
||: ||| |||  
Db 6 SEVKDAEFGXPKDD 20

RESULT 9  
US-08-660-531-18  
; Sequence 18, Application US/08660531  
; Patent No. 6221645  
; GENERAL INFORMATION:  
; APPLICANT: Chrysler, Susanna M.S.  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Keim, Pamela S.  
; APPLICANT: Anderson, John P.  
; TITLE OF INVENTION: Beta-Secretase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Ctr., 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA



ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/660,531  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/480,498  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Heslin, James M.  
REGISTRATION NUMBER: 29,541  
REFERENCE/DOCKET NUMBER: 15270-002210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-660-531-18

Query Match 22.8%; Score 33; DB 4; Length 21;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
II:::III I  
DB 1 SEVNLDAEFFRHD 12

RESULT 10  
US-08-313-200-8  
Sequence 8, Application US/08313200  
Patent No. 5998153  
GENERAL INFORMATION:  
APPLICANT: Baker, James R.  
APPLICANT: Koenig, Ronald J.  
TITLE OF INVENTION: THYROID PEROXIDASE EPITOPIC REGIONS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FERGSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/313,200  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kanski, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 20344-20658.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-313-200-8

Query Match 22.1%; Score 32; DB 2; Length 22;  
Best Local Similarity 54.5%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNK 24  
I:::III I  
DB 10 SRDITGAELEK 20

RESULT 11  
US-08-141-324-11  
Sequence 11, Application US/08141324  
Patent No. 5475097  
GENERAL INFORMATION:  
APPLICANT: Travis, James  
APPLICANT: Potempa, Jan S.  
APPLICANT: Barr, Philip J.  
APPLICANT: Pavloff, Nadine  
APPLICANT: Pike, Robert N.  
TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
TITLE OF INVENTION: Protease  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,324  
FILING DATE: 21-OCT-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 44-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-141-324-11

Query Match 22.1%; Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEFNKD 25  
II:::III I  
DB 4 SEVKKDAEFFRHD 15

RESULT 12  
US-08-141-324-12  
; Sequence 12, Application US/08141324  
; Patent No. 5475097  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; APPLICANT: Pike, Robert N.  
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 21-OCT-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 44-93  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303-499-8080  
; TELEFAX: 303-499-8089  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-08-141-324-12

Query Match 22.1%, Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEPFRD 25  
DB 4 SEVDLDAEPRRD 15

RESULT 13  
US-08-541-902-11  
; Sequence 11, Application US/08541902  
; Patent No. 5707620  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; APPLICANT: Pike, Robert N.  
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 21-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,324  
FILING DATE: 21-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 44-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-08-541-902-11

Query Match 22.1%, Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 SEISVGAEPFRD 25  
DB 4 SEVKMDAEPFRD 15

RESULT 14  
US-08-541-902-12  
; Sequence 12, Application US/08541902  
; Patent No. 5707620  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; APPLICANT: Pike, Robert N.  
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/541,902

FILING DATE: 435  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,324  
FILING DATE: 21-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Feider, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 44-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-541-902-12

Query Match 22.1%; Score 32; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 14 SEISVGAEFNKD 25  
||:|||||  
Db 4 SEVDLDEFRHD 15

RESULT 15  
US-08-324-301-3  
; Sequence 3, Application US/08324301  
; Patent No. 5597569  
; GENERAL INFORMATION:  
; APPLICANT: Siegall, Clay B.  
; APPLICANT: Gawlak, Susan L.  
; APPLICANT: Marguardt, Hans  
; TITLE OF INVENTION: A NEW RIBOSOME-INACTIVATING PROTEIN  
; TITLE OF INVENTION: ISOLATED FROM THE PLANT BRYONICA DIOLOCA  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bristol-Myers Squibb Company  
; STREET: 3005 First Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/324,301  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,891  
; FILING DATE: 25-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Poor, Brian W.  
; REGISTRATION NUMBER: 32,928  
; REFERENCE/DOCKET NUMBER: ONO109A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-728-4800  
; TELEFAX: 206-727-3601  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Bryonica dioica  
TISSUE TYPE: Root  
US-08-324-301-3

Query Match 21.4%; Score 31; DB 1; Length 19;  
Best Local Similarity 60.0%; Pred. No. 74;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 8 DVSLAFSEIS 17  
::|||||  
Db 2 NIELGFSEIS 11

RESULT 16  
US-08-706-741B-44  
; Sequence 44, Application US/08706741B  
; Patent No. 5955593  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: B3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAVERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63146  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/706,741B  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 965017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-706-741B-44

Query Match 20.7%; Score 30; DB 2; Length 13;  
Best Local Similarity 55.6%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CIGYHLDVS 10  
|:|:||||  
Db 5 CIGDEMDVS 13

RESULT 17  
US-08-924-695A-44  
; Sequence 44, Application US/08924695A  
; Patent No. 5998583

GENERAL INFORMATION:  
APPLICANT: KORSMEYER, STANLEY J.  
TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFFERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/924,695A  
FILING DATE: 09-SEP-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971798  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-924-695A-44

Query Match 20.7%; Score 30; DB 2; Length 13;  
Best Local Similarity 55.6%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHLDVS 10  
|:|:|  
Db 5 CIGDEMVS 13

RESULT 18  
US-08-440-391-7  
Sequence 7, Application US/08440391  
Patent No. 5656725  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
APPLICANT: LUTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,391  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-440-391-7

Query Match 20.7%; Score 30; DB 1; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHLDVS 10  
|:|:|  
Db 7 CIGDEMVS 15

RESULT 19  
US-08-440-391-32  
Sequence 32, Application US/08440391  
Patent No. 5656725  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
APPLICANT: LUTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,391  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-440-391-32

Query Match 20.7%; Score 30; DB 1; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CLGYHLDVS 10  
|:|:|  
Db 7 CIGDEMVS 15

RESULT 20  
US-08-248-839C-152  
; Sequence 152, Application US/08248839C  
; Patent No. 5843702  
; GENERAL INFORMATION:  
; APPLICANT: McConnell, David  
; APPLICANT: Devine, Kevin  
; APPLICANT: O'Kane, Charles  
; TITLE OF INVENTION: A Gene Expression System  
; NUMBER OF SEQUENCES: 185  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 58437020 No. 5843702disk of No. 5843702th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248,839C  
; FILING DATE: 25-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gregg, Valeta A.  
; REGISTRATION NUMBER: 35,127  
; REFERENCE/DOCKET NUMBER: 3614.214-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 152:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
; US-08-248-839C-152

Query Match 20.7%; Score 30; DB 2; Length 15;  
Best Local Similarity 46.2%; Pred. No. 80;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 10 SLAFSEISVGAEF 22  
DB 3 SLAFPAVICSSEF 15

RESULT 21  
US-08-908-597A-7  
; Sequence 7, Application US/08908597A  
; Patent No. 5863795  
; GENERAL INFORMATION:  
; APPLICANT: CHITTENDEN, Thomas D.; and  
; APPLICANT: LUTZ, Robert J.  
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/908,597A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,391  
; FILING DATE: 12-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, HENRY N.  
; REGISTRATION NUMBER: 32,073  
; REFERENCE/DOCKET NUMBER: 104322.147  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-942-8400  
; TELEFAX: 202-942-8484  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-908-597A-7

Query Match 20.7%; Score 30; DB 2; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CIGYHLDVS 10  
DB 7 CIGDEMDS 15

RESULT 22  
US-08-908-597A-32  
; Sequence 32, Application US/08908597A  
; Patent No. 5863795  
; GENERAL INFORMATION:  
; APPLICANT: CHITTENDEN, Thomas D.; and  
; APPLICANT: LUTZ, Robert J.  
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/908,597A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,391  
; FILING DATE: 12-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, HENRY N.  
; REGISTRATION NUMBER: 32,073  
; REFERENCE/DOCKET NUMBER: 104322.147  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-942-8400  
; TELEFAX: 202-942-8484  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-908-597A-32

## Query Match

Best Local Similarity 20.7%; Score 30; DB 2; Length 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 2 CIGYHLDVS 10  
1:1:111  
Db 7 CIGDEMDS 15

## RESULT 23

US-09-236-385A-7

Sequence 32, Application US/09236385A  
Patent No. 6221615

## GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/236,385A

FILING DATE: 25-Jan-1999

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073

TELECOMMUNICATION INFORMATION: (C) ATTORNEY DOCKET NO. 104322.147CIP

TELEPHONE: 202-942-8400

TELEFAX: 202-942-8484

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-236-385A-7

## Query Match

Best Local Similarity 20.7%; Score 30; DB 4; Length 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 2 CIGYHLDVS 10  
1:1:111  
Db 7 CIGDEMDS 15

## RESULT 24

US-09-236-385A-32

Sequence 32, Application US/09236385A  
Patent No. 6221615

## GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 41

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/236,385A

FILING DATE: 25-Jan-1999

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073

TELECOMMUNICATION INFORMATION: (C) ATTORNEY DOCKET NO. 104322.147CIP

TELEPHONE: 202-942-8400

TELEFAX: 202-942-8484

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-236-385A-32

## Query Match

Best Local Similarity 20.7%; Score 30; DB 4; Length 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;QY 2 CIGYHLDVS 10  
1:1:111  
Db 7 CIGDEMDS 15

## RESULT 25

PCT-US96-06122-7

Sequence 7, Application PC/TUS9606122

## GENERAL INFORMATION:

APPLICANT: IMMUNOGEN, INC.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS

WHICH MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/06122

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/440,391

FILING DATE: 12-MAY-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: WIXON, HENRY N.

REGISTRATION NUMBER: 32,073

REFERENCE/DOCKET NUMBER: 104322.147PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US96-06122-7

Query Match 20.7%; Score 30; DB 5; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 CIGYHLDVS 10  
I:I:III  
Db 7 CIGDEMDVS 15

Search completed: January 6, 2002, 09:13:06  
Job time: 424 sec

**THIS PAGE BLANK (USPTO)**